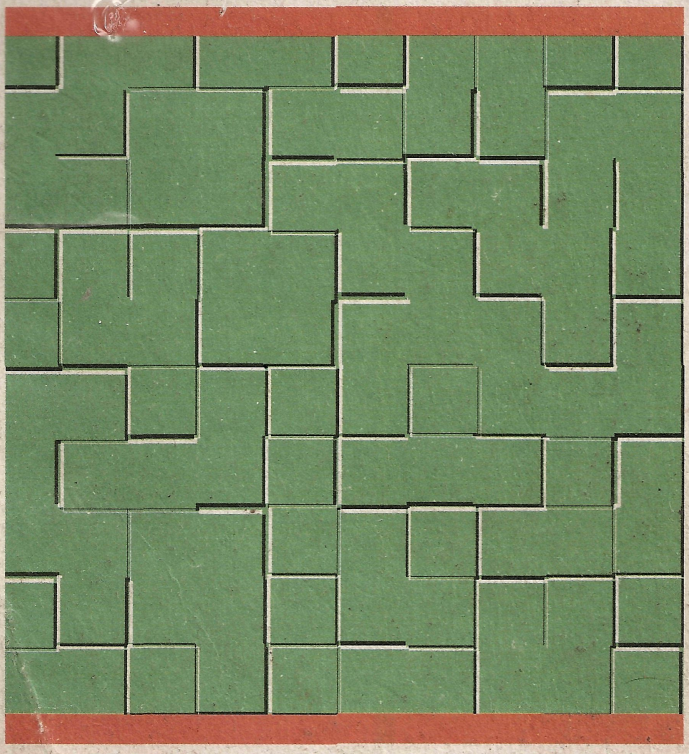
MANUALS

Matemàtiques



Manuel Castellet
Irene Llerena

Àlgebra lineal i geometria

Universitat Autònoma de Barcelona
Servei de Publicacions

4a EDICIÓ

MANUALS

Matemàtiques

Manuel Castellet és professor de Matemàtiques a la UAB des del 1976. Treballa en topologia algebraica. És director del Centre de Recerca Matemàtica de l'Institut d'Estudis Catalans.

Irene Llerena és professora de Matemàtiques a la UB des del 1976. Treballa en topologia algebraica. Ha fet l'assignatura Geometria I, base del contingut d'aquest llibre, durant els darrers set anys.

L'obra

Constitueix una exposició actual de la geometria afí i euclidiana des del punt de vista de l'àlgebra lineal.

És estructurada en 13 capítols. Al final de cada capítol s'inclou una breu nota històrica i una col·lecció d'exercicis de programar que han de permetre a l'estudiant aprofundir en la teoria i treballar mètodes de càlcul.

El llibre és especialment indicat per a un primer curs d'àlgebra lineal i/o geometria de qualsevol llicenciatura, enginyeria o diplomatura, especialment les científiques i tècniques.

En aquesta nova edició han estat millorades, corregides i ampliades algunes parts dels diferents capítols.

ISBN 84-7488-943-X

9 788474 889437

Universitat Autònoma de Barcelona
Servei de Publicacions

Índex

I Divisibilitat en els nombres enters

I.1 Divisió entera. Ideals 9
I.2 Mínim comú múltiple i màxim comú divisor 10
I.3 Nombres primers entre ells i nombres primers 12
I.4 Congruències 14
I.5 Els anells $\mathbf{Z}/(m)$ 16
I.6 Equacions diofàntiques lineals 17
I.7 Nota històrica 18
I.8 Exercicis 19
I.9 Exercicis de programar 20

II Divisibilitat a l'anell de polinomis

II.1 Definició de l'anell de polinomis 21
II.2 Divisió entera i ideals a $K[x]$ 23
II.3 Mínim comú múltiple i màxim comú divisor 25
II.4 Polinomis irreductibles i polinomis primers entre ells 27
II.5 Zeros d'un polinomi 29
II.6 Polinomis irreductibles de $\mathbf{R}[x]$ 31
II.7 Els anells $K[x]/(m(x))$ 32
II.8 Nota històrica 35
II.9 Exercicis 35
II.10 Exercicis de programar 36

III Grups

III.1 Definició i exemples 37
III.2 Permutacions 38
III.3 Subgrups 43
III.4 Homomorfismes 45
III.5 Grup quocient. Subgrups normals 46
III.6 Producte directe de grups 49
III.7 Grups cíclics 51

III.8 Grups finits 52
III.9 Nota històrica 56
III.10 Exercicis 57
III.11 Exercicis de programar 59

IV Espais vectorials

IV.1 Definició i exemples 61
IV.2 Subespais vectorials 63
IV.3 Bases d'un espai vectorial 64
IV.4 Fórmula de Grassmann. Suma directa de subespais. 69
IV.5 Suma directa d'espais vectorials 72
IV.6 Espai vectorial quocient. 73
IV.7 Coordenades 74
IV.8 Nota històrica 76
IV.9 Exercicis 77
IV.10 Exercicis de programar 79

V Aplicacions lineals

V.1 Definició i exemples 81
V.2 Matriu associada a una aplicació lineal 86
V.3 Teorema d'isomorfisme 90
V.4 L'espai de les aplicacions lineals 93
V.5 L'àlgebra d'endomorfismes 94
V.6 L'espai dual 96
V.7 Subespais ortogonals 99
V.8 Nota històrica 101
V.9 Exercicis 102
V.10 Exercicis de programar 104

VI Determinants

VI.1 Determinant de n vectors 105
VI.2 Determinant d'una matriu 110
VI.3 Determinant d'un endomorfisme 111
VI.4 Regla de Laplace 114
VI.5 Càlcul del rang d'una matriu 117
VI.6 Nota històrica 120
VI.7 Exercicis 121
VI.8 Exercicis de programar 122

VII Sistemes d'equacions lineals

VII.1 Planteig del problema 125
VII.2 Existència de solucions 126
VII.3 Regla de Cramer 127
VII.4 Resolució d'un sistema d'equacions lineals 127
VII.5 Mètode de Gauss 129
VII.6 Càlcul de la matriu inversa 133
VII.7 Nota històrica 133
VII.8 Exercicis 134
VII.9 Exercicis de programar 135

VIII Estructura dels endomorfismes

VIII.1 Vectors propis i valors propis. Polinomi característic 137
VIII.2 Diagonalització de matrius 140
VIII.3 Polinomi mínim 144
VIII.4 Subespais invariants 146
VIII.5 Grau del polinomi mínim 152
VIII.6 El teorema de Cayley-Hamilton 152
VIII.7 Matriu canònica (general) d'un endomorfisme 154
VIII.8 Matriu canònica de Jordan 158
VIII.9 Nota històrica 162
VIII.10 Exercicis 162
VIII.11 Exercicis de programar 165

IX Espais afins

IX.1 Definició d'espai afí 168
IX.2 Translacions. Una altra definició d'espai afí 169
IX.3 Varietats lineals 171
IX.4 Intersecció i suma de varietats lineals 173
IX.5 Dependència lineal de punts 175
IX.6 Coordenades baricèntriques 177
IX.7 Equacions d'una varietat en coordenades baricèntriques 182
IX.8 Coordenades cartesianes 183
IX.9 Equacions d'una varietat en coordenades cartesianes 185
IX.10 Raó simple 187
IX.11 Orientació d'un espai afí real 190
IX.12 Semiespais 191
IX.13 Nota històrica 192
IX.14 Exercicis 193
IX.15 Exercicis de programar 196

X **Afinitats**

X.1 Definició i primeres propietats 197
X.2 Uns exemples 200
X.3 Més propietats de les afinitats 204
X.4 Equacions d'una afinitat en una referència cartesiana 208
X.5 El grup afí 211
X.6 Varietats invariants 214
X.7 Classificació de les afinitats d'un espai afí A en ell mateix 216
X.8 Afinitats de la recta afí 217
X.9 Afinitats del pla afí 218
X.10 Nota històrica 222
X.11 Exercicis 222
X.12 Exercicis de programar 224

XI **Espais vectorials euclidians i unitaris**

XI.1 Formes bilineals i sesquilineals 225
XI.2 Producte escalar 227
XI.3 Norma 231
XI.4 Producte escalar i espai dual 233
XI.5 Subespais ortogonals 234
XI.6 Aplicacions adjuntes i autoadjuntes 235
XI.7 Diagonalització de matrius simètriques i hermítiques 236
XI.8 Producte vectorial 237
XI.9 Nota històrica 240
XI.10 Exercicis 240
XI.11 Exercicis de programar 242

XII **Aplicacions ortogonals. Aplicacions unitàries**

XII.1 Definicions 245
XII.2 Diagonalització de matrius unitàries 247
XII.3 Forma canònica d'una matriu ortogonal 248
XII.4 Els grups $O(2)$ i $SO(2)$ 250
XII.5 Angles 253
XII.6 El grup $O(3)$ 258
XII.7 Una altra determinació de les rotacions 261
XII.8 Composició de rotacions 261
XII.9 Nota històrica 265
XII.10 Exercicis 265
XII.11 Exercicis de programar 266

XIII Espais afins euclidians

XIII.1 Espais afins euclidians 269
XIII.2 Distància entre dues varietats lineals 271
XIII.3 Isometries 273
XIII.4 Classificació dels desplaçaments 275
XIII.5 Desplaçaments de la recta euclidiana 275
XIII.6 Desplaçaments del pla euclidià 276
XIII.7 Desplaçaments de l'espai euclidià tridimensional 278
XIII.8 Semblances 281
XIII.9 Semblances de l'espai afí euclidià tridimensional 283
XIII.10 Semblances del pla afí euclidià 283
XIII.11 Alguns exemples i aplicacions 285
XIII.12 Nota històrica 292
XIII.13 Exercicis 292
XIII.14 Exercicis de programar 296

Capítol I

Divisibilitat en els nombres enters

I.1 Divisió entera. Ideals

Designarem per $\mathbf{Z}$ el conjunt dels nombres enters. La teoria de la divisibilitat a $\mathbf{Z}$ és conseqüència de la següent important propietat.

Teorema 1.1 (de la divisió entera) *Donats $a, b \in \mathbf{Z}$, $b \neq 0$, existeixen dos únics nombres enters q i r que compleixen $a = b \cdot q + r$, $0 \leq r < |b|$. q i r es diuen* el quocient *i* la resta *de la divisió entera de a per b.*

Exemple:

$$-8 = 3 \cdot (-3) + 1, \qquad 3 = (-8) \cdot 0 + 3.$$

Si la resta de la divisió entera de a per b és 0, es diu que a és un *múltiple* de b (escriurem $a = \dot{b}$), que b és un *divisor* de a (escriurem $b \mid a$), o que a és *divisible* per b. Indicarem per (b) el conjunt dels múltiples de b. Observem que (b) compleix les dues propietats següents:

- és tancat per la suma; és a dir $a, c \in (b) \Rightarrow a + c \in (b)$.
- si $a \in (b)$ i c és qualsevol enter, aleshores $a \cdot c \in (b)$.

Proposició 1.2 *Si el subconjunt $I \subset \mathbf{Z}$ compleix*

1. $a, b \in I \Rightarrow a + b \in I$
2. $a \in I$, $c \in \mathbf{Z} \Rightarrow a \cdot c \in I$,

aleshores hi ha un $b \in I$ tal que $I = (b)$.

DEMOSTRACIÓ: Si $I = \{0\}$, aleshores $I = (0)$. Si I conté un element no nul a, també conté $-a = a \cdot (-1)$, i o bé a o bé $-a$ és positiu. Per tant, I conté enters positius. Sigui b el més petit dels positius continguts a I. Per 2. I conté tots els múltiples de b: $(b) \subset I$. Anem a veure que $I \subset (b)$ i per tant $I = (b)$. En efecte, donat $a \in I$ qualsevol, per (1.1)

$$a = b \cdot q + r.$$

Per 1 i 2 $r = a - b \cdot q = a + b \cdot (-q) \in I$; però $0 \leq r < |b| = b$, i b és el més petit dels positius de I; cal doncs que $r = 0$ i , per tant, $a = b \cdot q \in (b)$. □

Un subconjunt I que compleixi 1 i 2 com a (1.2) es diu un *ideal* de $\mathbf{Z}$. L'element b tal que $I = (b)$ es diu *base* de l'ideal.

Exercici:

$(b) = (c)$ si i només si $c = \pm b$.

Observació:

$(a) \subset (b)$ si i només si $b \mid a$. Les qüestions de divisibilitat equivalen, per tant, a qüestions sobre inclusions entre ideals.

I.2 Mínim comú múltiple i màxim comú divisor

Donats nombres enters $a_1, \ldots, a_n$, la intersecció $(a_1) \cap \ldots \cap (a_n)$ és el conjunt dels nombres enters múltiples comuns de tots ells. Aquest conjunt compleix les dues condicions de (1.2) i, per tant, $(a_1) \cap \ldots \cap (a_n) = (m)$ per a un m convenient. Aquest m està caracteritzat per les dues propietats següents:

- m és múltiple comú de $a_1, \ldots, a_n$.
- qualsevol altre múltiple comú de $a_1, \ldots, a_n$ és múltiple de m.

Direm que m és el *mínim comú múltiple de* $a_1, \ldots, a_n$ i escriurem

$$m = \text{m.c.m.}(a_1, \ldots, a_n).$$

Atenció!:

Observem que també $-m$ és mínim comú múltiple de $a_1, \ldots, a_n$.

Considerem ara la unió $(a_1) \cup \ldots \cup (a_n)$. Aquest conjunt, en general, no compleix les condicions de (1.2). Per exemple $(2) \cup (3)$ no conté el $5 = 2 + 3$. Formem a partir de $(a_1) \cup \ldots \cup (a_n)$ un subconjunt I de $\mathbf{Z}$ que compleixi les condicions de (1.2). Per

1. I ha de contenir totes les sumes de múltiples de $a_1, \ldots, a_n$: $a_1c_1 + \ldots + a_nc_n$. No cal ampliar més; el conjunt

$$I = \{a_1c_1 + \ldots + a_nc_n \mid c_1, \ldots, c_n \in \mathbf{Z}\}$$

compleix ja les condicions de (1.2) i, per tant, existeix un enter d tal que $I = (d)$. Denotarem I per $(a_1, \ldots, a_n)$. Així $I = (a_1, \ldots, a_n) = (d)$. d està caracteritzat per les dues propietats següents:

- d és divisor comú de $a_1, \ldots, a_n$, ja que això equival a dir que $a_i \in (d)$ per a $i = 1, \ldots, n$. $(a_i = a_1 \cdot 0 + \ldots + a_i \cdot 1 + \ldots + a_n \cdot 0 \in I)$.
- Tot altre divisor d' comú a $a_1, \ldots, a_n$ divideix d. En efecte, que d' sigui divisor de $a_1, \ldots, a_n$ vol dir que $a_i \in (d')$, $i = 1, \ldots, n$. Per tant, $\{a_1c_1 + \ldots + a_nc_n \mid c_i \in \mathbf{Z}\} \subset (d')$, és a dir, $(d) \subset (d')$, la qual cosa vol dir que d' és un divisor de d.

Direm que d és el *màxim comú divisor de* $a_1, \ldots, a_n$ i escriurem

$$d = \text{m.c.d.}(a_1, \ldots, a_n).$$

Atenció!:

També $-d$ és màxim comú divisor de $a_1, \ldots, a_n$.

Observem que el màxim comú divisor d és una suma de múltiples de $a_1, \ldots, a_n$,

$$d = a_1 \cdot r_1 + \ldots + a_n \cdot r_n.$$

Aquesta expressió es coneix amb el nom d'*identitat de Bézout* .

Acabarem aquest apartat amb un mètode pràctic de càlcul del màxim comú divisor i de la identitat de Bézout. El mètode es basa en el següent resultat:

Proposició 2.1 *Sigui* $a = b \cdot q + r$ *la divisió entera de* a *per* b. *Aleshores*

$$\text{m.c.d.}(a, b) = \text{m.c.d.}(b, r).$$

DEMOSTRACIÓ: El resultat és conseqüència del fet que $(a, b) = (b, r)$. En efecte, tot element $ac_1 + bc_2 \in (a, b)$ és $ac_1 + bc_2 = b(qc_1 + c_2) + rc_1 \in (b, r)$ i, recíprocament, tot element $bn_1 + rn_2 \in (b, r)$ és $bn_1 + rn_2 = an_2 + b(n_1 - qn_2) \in (a, b)$. □

Si apliquem reiteradament aquesta proposició, obtenim

$$\begin{array}{lll} a = b \cdot q + r & (a, b) = (b, r) & r < |b| \\ b = r \cdot q_1 + r_1 & (b, r) = (r, r_1) & r_1 < r \\ r = r_1 \cdot q_2 + r_2 & (r, r_1) = (r_1, r_2) & r_2 < r_1. \end{array}$$

Les restes successives van disminuint i obtindrem, per tant, en un moment donat, la resta zero:

$$\begin{array}{ll} r_{k-2} = r_{k-1} \cdot q_k + r_k & (r_{k-2}, r_{k-1}) = (r_{k-1}, r_k) \qquad r_k < r_{k-1} \\ r_{k-1} = r_k \cdot q_{k+1} + 0 & (r_{k-1}, r_k) = (r_k, 0) = (r_k). \end{array}$$

Així, doncs, $(a, b) = (r_k)$; és a dir, $r_k = \text{m.c.d.}(a, b)$.

Aquest mètode de trobar el màxim comú divisor es diu *algorisme d'Euclides* .

Per calcular el màxim comú divisor de més de dos enters apliquem:

Exercici:

$\text{m.c.d.}(a_1, a_2, a_3) = \text{m.c.d.}[\text{m.c.d.}(a_1, a_2), a_3]$ i, en general,
$\text{m.c.d.}(a_1, \ldots, a_n) = \text{m.c.d.}[\text{m.c.d.}(a_1, \ldots, a_{n-1}), a_n]$.

Les divisions enteres efectuades en l'algorisme d'Euclides ens permeten expressar $d = r_k = \text{m.c.d.}(a, b)$ com a suma d'un múltiple de a i un múltiple de b. En efecte, a

$$d = r_k = r_{k-2} - r_{k-1} \cdot q_k$$

d s'expressa com a suma d'un múltiple de r_{k-2} i un múltiple de r_{k-1}. Ara bé, $r_{k-1} = r_{k-3} - r_{k-2} \cdot q_{k-1}$, i substituint a la igualtat anterior obtenim una expressió de d com a suma d'un múltiple de r_{k-3} i un múltiple de r_{k-2}. Tornant a substituir convenientment podem expressar d com a suma de múltiples de r_{k-4} i r_{k-3}; i així successivament fins a obtenir la identitat de Bézout

$$d = a \cdot r + b \cdot s.$$

En el pròxim apartat (3.2) provarem que si $m = \text{m.c.m.}(a, b)$ i $d = \text{m.c.d.}(a, b)$, $m \cdot d = \pm a \cdot b$ Això ens permet calcular m si coneixem d. Per al càlcul del mínim comú múltiple de més de dos nombres fem servir:

Exercici:

$\text{m.c.m.}(a_1, a_2, a_3) = \text{m.c.m.}[\text{m.c.m.}(a_1, a_2), a_3]$ i, en general,
$\text{m.c.m.}(a_1, \ldots, a_n) = \text{m.c.m.}[\text{m.c.m.}(a_1, \ldots, a_{n-1}), a_n]$.

I.3 Nombres primers entre ells i nombres primers

Es diu que a i b són *primers entre ells* si $\text{m.c.d.}(a, b) = 1$.

Exemples:

1. $\text{m.c.d.}(3, 8) = 1$. Observem que $1 = 3 \cdot 3 + 8(-1)$.
2. Si $d = \text{m.c.d.}(a, b)$ i $a = d \cdot a'$, $b = d \cdot b'$, aleshores $\text{m.c.d.}(a', b') = 1$. En efecte, si d' fos un divisor comú de a' i b', $d \cdot d'$ seria divisor comú de a i b i, per tant, un divisor de d. Això només és possible si $d' = \pm 1$.

Teorema 3.1 (d'Euclides) *Si* $a \mid b \cdot c$ *i* $\text{m.c.d.}(a,b) = 1$, *aleshores* $a \mid c$.

DEMOSTRACIÓ: Si $1 = \text{m.c.d.}(a,b)$ podem expressar l'1 com $1 = a \cdot r + b \cdot s$. Multiplicant per c obtenim $c = acr + bcs$. a divideix els dos sumands i, per tant, $a \mid c$. □

Proposició 3.2 *Si* $m = \text{m.c.m.}(a,b)$ *i* $d = \text{m.c.d.}(a,b)$, *es compleix* $m \cdot d = \pm a \cdot b$.

DEMOSTRACIÓ: Posem $a = da'$ i $b = db'$. Es tracta de veure que $m = \pm da'b'$ és un mínim comú múltiple de a i b. Que $da'b'$ és múltiple comú de a i b és evident. Sigui n un altre múltiple comú de a i b ; és a dir, $n = ar = bs$. Llavors $a'dr = b'ds$, d'on $a'r = b's$ amb a', b' primers entre ells. Aleshores, per (3.1), a' divideix s, és a dir, $s = a'h$ i $n = bs = db'a'h$. Així resulta que n és múltiple de $db'a'$. □

Qualsevol nombre enter p és divisible per ± 1 i per $\pm p$. Direm que p és *primer* si aquests són els seus únics divisors. L'1 i el -1 no es consideren nombres primers.

Proposició 3.3 *El conjunt dels nombres primers és infinit.*

DEMOSTRACIÓ: Ho demostrarem veient que, donat un conjunt finit de nombres primers $N = \{p_1, \ldots, p_m\}$, sempre hi ha un nombre primer fora de N. En efecte, considerem $a = p_1 \cdots p_m + 1$. Si $b \mid a$, també $-b \mid a$; per tant, a té sempre divisors positius. Sigui p el més petit dels divisors positius de a diferents de 1. Clarament p és primer. Si p fos un dels p_i, dividiria $p_1 \cdots p_m$ i, per tant, dividiria $a - p_1 \cdots p_m = 1$. Això és impossible , ja que $p \neq 1$. D'aquí que $p \notin N$. □

Proposició 3.4 *Tot nombre enter* a *no nul,* $a \neq \pm 1$, *és producte de nombres primers.*

DEMOSTRACIÓ: Com hem vist a la demostració de (3.3) a té sempre un divisor primer $p_1 \neq \pm 1$. Tenim així $a = p_1 \cdot a_1$. Si $a_1 \neq \pm 1$, escollim un divisor primer de a_1, $p_2 \neq \pm 1$, i tindrem $a_1 = p_2 \cdot a_2$. Així $a = p_1 \cdot p_2 \cdot a_2$. Repetim el mateix procés si $a_2 \neq \pm 1$, i així successivament. Ara bé, $|a| > |a_1| > |a_2| > \ldots$. Arribarà, doncs, un moment en què tindrem $a = p_1 \cdots p_n a_n$ amb $a_n = \pm 1$. Això és una descomposició en nombres primers de a. □

La descomposició d'un enter en producte de primers no és ben bé única. Per exemple,

$$12 = 2 \cdot 2 \cdot 3 = 2 \cdot (-2) \cdot (-3) = (-2) \cdot 3 \cdot (-2).$$

Hi ha, però, una certa unicitat. Concretament,

Proposició 3.5 *Si* $p_1 \cdots p_n = q_1 \cdots q_m$ *i tots els factors* p_i, q_j, *amb* $i = 1, \ldots, n$, $j = 1, \ldots, m$, *són nombres primers, aleshores* $n = m$ *i els nombres* $\{p_1, \ldots, p_n\}$ *són els mateixos que els* $\{q_1, \ldots, q_m\}$, *llevat del signe (i l'ordre).*

DEMOSTRACIÓ: Observem que si p,q són nombres primers, o bé m.c.d.$(p,q) = 1$ o bé $p = \pm q$. p_1 divideix $p_1 \cdots p_n = q_1(q_2 \cdots q_m)$. Pel teorema d'Euclides (3.1), o bé $p_1 \mid q_2 \cdots q_m$, quan m.c.d.$(p_1, q_1) = 1$, o bé $p_1 = \pm q_1$. En el primer cas $p_1 \mid q_2(q_3 \cdots q_m)$; aplicant de nou el teorema d'Euclides obtenim que $p_1 \mid q_3 \cdots q_m$, o $p_1 = \pm q_2$. Repetim el procés tantes vegades com sigui necessari. O bé trobarem que p_1 és un dels $q_j, j = 1, \ldots, m-2$, llevat del signe, o bé conclourem que $p_1 \mid q_{m-1} \cdot q_m$. D'on $p_1 = \pm q_{m-1}$ o $p_1 = \pm q_m$.

Així doncs, p_1 coincideix, llevat del signe, amb un dels q_j. Canviant l'ordre si convé, podem suposar que $p_1 = \pm q_1$. Llavors $p_2 \cdots p_n = (\pm q_2) \cdot q_3 \cdots q_m$. El mateix raonament prova que p_2 és igual, llevat del signe, a un dels q_j, $j = 2, \ldots, m$. I així successivament. Si $n < m$, arribarem a la situació $1 = \pm q_{n+1} \cdots q_m$, i això no és possible perquè tots els q_j són diferents de ± 1. Si $m > n$, arribarem a $\pm p_{m+1} \cdots p_n = 1$, igualment impossible. Per tant, $n = m$. □

I.4 Congruències

Fixem $0 \neq m \in \mathbf{Z}$. Direm que dos nombres enters a i b són *congruents mòdul* m si $a - b \in (m)$. Això equival a dir que les divisions enteres de a i b per m tenen la mateixa resta. En efecte,

$$\left.\begin{array}{rcl} a & = & m \cdot q + r \\ b & = & m \cdot q_1 + r_1 \end{array}\right\} \Rightarrow a - b = m(q - q_1) + (r - r_1) \text{ amb } |r - r_1| < |m|.$$

Per tant, $a - b \in (m)$ si i només si $r = r_1$. Si a i b són congruents mòdul m escriurem $a \equiv b\,(m)$.

És molt fàcil veure que es compleixen les següents condicions:

1. Per a tot $a \in \mathbf{Z}$, $a \equiv a\,(m)$.
2. $a \equiv b\,(m) \Rightarrow b \equiv a\,(m)$.
3. $a \equiv b\,(m)$, $b \equiv c\,(m) \Rightarrow a \equiv c\,(m)$.

Formem ara subconjunts de $\mathbf{Z}$ de la següent manera: cada subconjunt està format per tots els nombres enters que donen la mateixa resta en fer la divisió entera per m. Obtenim m subconjunts:

- $(m) =$ conjunt d'enters que donen la resta 0
- $\{\dot{m} + 1\} =$ conjunt d'enters que donen la resta 1
- .....
- $\{\dot{m} + (|m| - 1)\} =$ conjunt d'enters que donen la resta $|m| - 1$.

Aquests conjunts s'anomenen *classes de restes mòdul m*. Designarem per $\mathbf{Z}/(m)$ el conjunt de les classes de restes mòdul m. Cada enter està en una d'aquestes classes i només en una. Una classe queda, per tant, ben determinada donant un qualsevol dels seus elements. Direm que aquest element és un *representant* de la classe.

Nota:

El procés que acabem de fer és un cas particular d'un procés general molt usual a matemàtiques. Es tracta del següent: sigui A un conjunt. Una *relació* a A és un criteri que ens permet dir si dos elements qualssevol de A, a i b, "satisfan la relació" o no. Més exactament: donar una relació a A és donar una sèrie de parells ordenats d'elements de A (que seran els elements que "satisfan la relació"); és a dir, donar un subconjunt del producte cartesià $A \times A$. Indiquem per $a \sim b$ el fet que a estigui relacionat amb b. Exemples de relacions són

- $a \sim b \Leftrightarrow a \mid b$.
- $a \sim b \Leftrightarrow a < b$.
- $a \sim b \Leftrightarrow a - b \in (m)$.

Una relació és *relació d'equivalència* si compleix

- Propietat reflexiva. Per a tot $a \in A$, $a \sim a$.
- Propietat simètrica: $a \sim b \Rightarrow b \sim a$.
- Propietat transitiva: $a \sim b$, $b \sim c \Rightarrow a \sim c$

Dels exemples anteriors només la congruència mòdul m és una relació d'equivalència. Tota relació d'equivalència ens permet dividir el conjunt A en subconjunts disjunts (*classes d'equivalència*) de la següent manera: cada classe està formada per tots els elements relacionats entre ells. Les tres propietats anteriors asseguren que tot element és en una i només en una classe. En efecte, designem per $[a]$ la classe de tots els elements relacionats amb a. Clarament $a \in [a]$. Suposem que a és també en una altra classe: $a \in [c]$. Llavors $a \sim c$ i les propietats transitiva i simètrica ens diuen que tot element relacionat amb a està també relacionat amb c i viceversa. És a dir, $[a] = [c]$.

Una *partició* de A és una sèrie de subconjunts de A tals que tot $a \in A$ és en un i només en un d'aquests subconjunts. Una *classificació* dels elements de A no és altra cosa que una partició de A. Per exemple, classifiquem $\mathbf{Z}$ en parells i senars, o classifiquem les persones per la nacionalitat. Les classes d'equivalència formen una partició de A. Recíprocament, una partició de A determina una relació d'equivalència: a, $b \in A$ són "equivalents" si són en el mateix subconjunt de la partició. Per tant, classificar és el mateix que formar classes per una relació d'equivalència. Aquesta relació ve a ser el criteri pel qual classifiquem. Per exemple, si volem classificar els nombres

enters, haurem de fixar amb quin criteri ho fem. Si ho fem per la paritat, considerarem dos enters a la mateixa classe si tots dos són parells o tots dos són senars. El que hem fet no és sinó donar una relació d'equivalència.

El conjunt de les classes d'equivalència es diu *conjunt quocient* i es denota per $A/\sim$.

I.5 Els anells $\mathbf{Z}/(m)$

Volem ara definir unes operacions a $\mathbf{Z}/(m)$ que facin el paper que feien la suma i el producte a $\mathbf{Z}$. La manera més natural és definir

$$[a] + [b] = [a + b] \qquad [a] \cdot [b] = [ab].$$

Hi ha un problema, però. Considerem uns altres representants de les classes $[a]$ i $[b]$: siguin $[a_1] = [a], [b_1] = [b]$. Les mateixes definicions donen $[a_1] + [b_1] = [a_1 + b_1]$, $[a_1] \cdot [b_1] = [a_1 b_1]$. Les classes $[a_1 + b_1]$, $[a_1 b_1]$ que ara obtenim, coincideixen amb les classes $[a + b]$, $[ab]$ obtingudes abans? En altres paraules, la suma i el producte definits, depenen dels representants escollits? La resposta és no; en efecte,

$$\left.\begin{array}{lll} [a_1] = [a] & \Rightarrow & a_1 = a + \dot{m} \\ [b_1] = [b] & \Rightarrow & b_1 = b + \dot{m} \end{array}\right\} \Rightarrow \left\{\begin{array}{lll} a_1 + b_1 = a + b + \dot{m} & \Rightarrow & [a_1 + b_1] = [a + b] \\ a_1 b_1 = ab + \dot{m} & \Rightarrow & [a_1 b_1] = [ab]. \end{array}\right.$$

Un conjunt A amb dues operacions $(a + b,\ a \cdot b)$ és un *anell* si compleix

- Propietats de $+$:
  - Associativa: $(a + b) + c = a + (b + c) \quad \forall a, b, c \in A$
  - Commutativa: $a + b = b + a \quad \forall a, b \in A$
  - Existeix un element, que anomenarem *zero* i escriurem 0, tal que
    $$a + 0 = 0 + a = a \quad \forall a \in A$$
  - Per a cada $a \in A$ hi ha un element, que anomenarem *l'oposat* de a i denotarem $-a$, tal que $a + (-a) = 0$
- Propietat de $\cdot$:
  - Associativa: $(a \cdot b) \cdot c = a \cdot (b \cdot c) \quad \forall a, b, c \in A$
- Propietats que relacionen $+$ i $\cdot$:
  - Distributives:
    $$a \cdot (b + c) = a \cdot b + a \cdot c$$
    $$(a + b) \cdot c = a \cdot c + b \cdot c \quad \forall a, b, c \in A.$$

Si, a més a més, es compleix que l'operació $\cdot$ és commutativa ($a \cdot b = b \cdot a$, per a tot $a,b \in A$) es diu que A és un *anell commutatiu* . Si existeix un element $e \in A$ tal que $a \cdot e = e \cdot a = a$, per a tot $a \in A$, es diu que A *té unitat* . L'element e es diu *la unitat de* A i generalment es designa per 1. Un element $a^{-1} \in A$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$ es diu un *invers* de a.

Observem que en un anell $a \cdot 0 = 0 \cdot a = 0$ per a tot a. En efecte, $a \cdot 0 = a(0+0) = a \cdot 0 + a \cdot 0$. Per tant, sumant $-(a \cdot 0)$ als dos costats, obtenim $0 = a \cdot 0$. Resulta doncs que en un anell A el 0 no pot tenir mai invers. Un anell commutatiu amb unitat, en què tot element no zero té invers, es diu un *cos*. $\mathbf{Z}$ és un anell commutatiu amb unitat. El conjunt de racionals $\mathbf{Q}$, el conjunt de reals $\mathbf{R}$ i el conjunt de complexos $\mathbf{C}$ són un cos.

$\mathbf{Z}/(m)$ és un anell commutatiu amb unitat, [1]. $\mathbf{Z}/(m)$ té, però, propietats que no tenia $\mathbf{Z}$. Per exemple, el producte de dos elements diferents de [0] pot ser [0]. Així a $\mathbf{Z}/(6)$, $[2] \cdot [3] = [0]$. Aquests elements es diuen *divisors de zero* . Per altra banda hi ha elements que tenen invers. Per exemple, a $\mathbf{Z}/(8)$, $[3] \cdot [3] = [1]$. Observem que, en un anell, si un element és divisor de zero no pot tenir invers. En efecte, sigui $a \cdot b = 0$ amb $a \neq 0$ i $b \neq 0$. Si existeix l'invers de a, resulta $b = 1 \cdot b = (a^{-1} \cdot a) \cdot b = a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 = 0$, en contra del que hem suposat.

Proposició 5.1 *Si* m.c.d.$(a,m) = 1$, $[a]$ *té un invers a* $\mathbf{Z}/(m)$. *Si* m.c.d.$(a,m) = d \neq \pm 1, \pm m$, $[a]$ *és un divisor de zero a* $\mathbf{Z}/(m)$.

DEMOSTRACIÓ: Si m.c.d.$(a,m) = 1$ podem posar $1 = ar + ms$. D'on $[1] = [ar] = [a][r]$ i $[r]$ és invers de $[a]$. Si $d =$ m.c.d.(a,m) posem $a = d \cdot a'$, $m = d \cdot m'$. Llavors $am' = a'm \in [0]$, d'on $[a][m'] = [0]$ i $[m'] \neq 0$, ja que $0 < m' < |m|$.□

Corol·lari 5.2 *L'anell* $\mathbf{Z}/(p)$ *és un cos si i només si* p *és primer.*

DEMOSTRACIÓ: Si p és primer, (5.1) ens diu que $\mathbf{Z}/(p)$ és un cos. Si $\mathbf{Z}/(p)$ és un cos, no pot tenir divisors de zero (veure l'observació feta abans de (5.1)). Aleshores (5.1) ens diu que p ha d'ésser primer. □

I.6 Equacions diofàntiques lineals

El nostre objectiu en aquest apartat és estudiar les solucions enteres de l'equació

$$ax + by = c$$

on $a,b,c \in \mathbf{Z}$. La primera proposició es refereix a l'existència de solucions.

Proposició 6.1 *L'equació diofàntica* $ax + by = c$, $a,b,c \in \mathbf{Z}$, *té solució si i només si el màxim comú divisor de* a *i* b *divideix* c.

Exercici:

Demostreu aquesta proposició.

Suposem, doncs, que $ax + by = c$ té solució. Dividint per $d = \text{m.c.d.}(a, b)$ obtenim una equació amb les mateixes solucions, $a'x + b'y = c'$, en la qual $\text{m.c.d.}(a', b') = 1$. Multipliquem la identitat de Bézout $1 = a'r + b's$ per c':

$$c' = a'rc' + b'sc'.$$

$x = rc'$, $y = sc'$ és doncs una solució de l'equació $a'x + b'y = c'$.

Per altra banda, restant dues de les expressions anteriors obtenim

$$a'(x - rc') + b'(y - sc') = 0.$$

Pel Teorema d'Euclides (3.1)

$$a' \mid y - sc' \qquad \text{i} \qquad b' \mid x - rc'.$$

És a dir, existeixen t i u tals que

$$\begin{aligned} y &= sc' + ta' \\ x &= rc' + ub'. \end{aligned}$$

Substituint a l'equació inicial

$$c' = a'x + b'y = a'rc' + a'ub' + b'sc' + b'ta' = c' + a'b'(u + t),$$

ja que rc', sc' és una solució. Per tant, $u + t = 0$. La solució general de l'equació donada és, doncs,

$$\begin{aligned} x &= rc' - tb' \\ y &= sc' + ta'. \end{aligned}$$

I.7 Nota històrica

L'aritmètica, que s'inicià amb els babilònics cap a l'any 2000 a. C. i es desenvolupà entre els anys 600 i 300 a. C. a les escoles gregues de Pitàgores, Euclides i Diofant, és avui encara una branca d'intensa i atractiva activitat investigadora. Les propietats dels nombres enters i les relacions entre ells, els conceptes i propietats de múltiple, divisor, nombre primer, la descomposició d'un enter (positiu) en producte de primers, el teorema d'Euclides, etc. formaren ja part del cos de doctrina dels llibres VII, VIII i IX dels *Elements* d'Euclides. Pierre de Fermat (1601?–1665), un home de lletres que llegia matemàtiques per afició (l'*Aritmètica* de Diofant d'Alexandria) és una de les figures clau de l'aritmètica moderna; ell

fou qui es plantejà resoldre la majoria de problemes aritmètics donant alguns criteris, demostrant teoremes, establint conjectures i assegurant haver demostrat un resultat (conegut ara com el darrer teorema de Fermat) que, malgrat els esforços dels més il·lustres matemàtics, encara resta com una qüestió oberta: l'equació $x^n + y^n = z^n$ amb x, y, z enters i $n > 2$, no té cap solució no trivial. És el gran repte (o la gran espina) que tenen els investigadors en teoria de nombres. Sense passar per alt la contribució de Leonhard Euler (1707–1783) que, entre altres, va demostrar el 1736 el petit teorema de Fermat: $a^p \equiv a\,(p)$, p primer, i la de Carl Friedrich Gauss (1777–1855), que a les seves *Disquisitiones Arithmeticae* va sistematitzar les congruències i en desenvolupà la teoria tal com ho fem avui, convé esmentar Ernst Eduard Kummer (1810–1893), Julius Wilhelm Richard Dedekind (1831–1916) i Leopold Kronecker (1823–1891), els quals, en llurs treballs sobre nombres algebraics, utilitzen ja els conceptes d'anell, ideal i cos, encara que les teories abstractes no s'han desenvolupat fins el segle 20.

I.8 Exercicis

1. Calculeu m.c.d.$(28n + 5, 35n + 2)$ per a tot $n \geq 1$.

2. Proveu que a la successió de Fibonacci 0, 1, 1, 2, 3, 5, 8, 13, ... ($a_n = a_{n-1} + a_{n-2}$) dos termes consecutius són sempre primers entre ells.

3. Demostreu que si p és primer $(p-1)! \equiv -1\,(p)$ (congruència de Wilson).

4. Demostreu que si p és primer $a^p \equiv a\,(p)$ per a tot a (petit teorema de Fermat).

5. Calculeu 2001^{2001} mòdul 17.

6. Demostreu els criteris de divisibilitat per 3, 4, 5, 9, 11, 13 i 19.

7. Resoleu les equacions diofàntiques $111x + 36y = 15$, $10x + 26y = 1224$, $6x + 10y = 20$, $6x + 10y = 3$.

8. En una illa deserta —només habitada per un mono i molts cocoters— arriben cinc nàufregs; recullen tants cocos com poden i es posen a descansar. A mitja nit, un mariner desconfiat, temerós que els altres es despertin i es mengin algun coco, es lleva, fa cinc parts iguals del total de cocos, separa la seva part i deixa la resta; però li ha sobrat un coco que dóna al mono. Al cap d'una hora un segon mariner té la mateixa pensada: fa cinc parts iguals del total de cocos (dels que queden, és clar!), se'n guarda una part, deixa la resta i dóna al mono un coco que ha sobrat. Al cap d'una altra hora, ... Cada un dels cinc mariners fa la mateixa operació.

   L'endemà al matí en llevar-se decideixen repartir els cocos (els del piló final) entre els cinc (cada un d'ells rient per sota el nas). Sobra un coco que donen al mono. Pregunta: quants cocos havien collit? (The Saturday Evening Post, $\simeq$ 1925).

9. L'Oliana Molls treballa quatre dies seguits i en descansa un. La Betty en treballa dos i en descansa un. Només es veuen els dies de lluna plena (un de cada vint-i-vuit dies). La Betty va tenir lliure ahir, l'Oliana el tindrà demà passat i fa deu dies que era lluna plena. Quants dies falten perquè es vegin? Quants dies lliures comuns hauran perdut mentrestant per falta de lluna plena? (TV3, 1986).

10. a) Trobeu les solucions de l'equació lineal $6x \equiv 14\ (16)$, i de l'equació de segon grau $x^2 - 3x - 3 \equiv 0\ (7)$.

b) Estudieu, en general, la resolució de $ax \equiv b\ (m)$, $ax^2 + bx + c \equiv 0 (p)$ amb p primer.

11. (Teorema xinès de la resta) Demostreu que si $(m, n) = 1$ les equacions $x \equiv a\,(m)$ i $x \equiv b\,(n)$ tenen una única solució mòdul mn.

12. Determineu els $a \in \mathbf{Z}/(8)$ tals que el sistema $7x + 5y = 2$, $5x + ay = 16$ té solució a $\mathbf{Z}/(8)$.

13. a) Demostreu que si $(a, n) = (b, n) = 1$ l'equació $ax + by = c$ té exactament n solucions a $\mathbf{Z}/(n)$.

b) Trobeu les solucions de $3x + 4y = 1$ a $\mathbf{Z}/(7)$ i de $3x + 7y = 2$ a $\mathbf{Z}/(8)$.

14. Demostreu que en qualsevol solució entera x,y,z de l'equació $x^2 + y^2 = z^2$ *(terna pitagòrica)*,

a) x, y o z és múltiple de 5

b) x o y és múltiple de 3

c) x o y és múltiple de 4.

15. Demostreu que les úniques relacions d'equivalència a $\mathbf{Z}$ compatibles amb la suma i el producte són les congruències.

I.9 Exercicis de programar

16. Càlcul del màxim comú divisor i del mínim comú múltiple de dos nombres enters. (Indicació: utilitzeu les proposicions I.2.1 i I.3.2).

17. Resolució de l'equació diofàntica $ax + by = c$. (Indicació: utilitzeu com a subrutina l'exercici I.16 i seguiu el procés de l'apartat I.6).

18. Factorització d'un nombre enter en producte de primers.

19. Construcció de la taula dels nombres primers més petits que 100.000. (Indicació: aneu guardant els primers més petits o iguals que 313 en una variable dimensionada. Així els tindreu disponibles per anar fent les successives divisions).

20. Càlcul de $1/a$ a $\mathbf{Z}/(p)$, $a \neq 0$, p primer.

21. Càlcul de $\sqrt{a}$ a $\mathbf{Z}/(p)$, p primer, si existeix.

Capítol II

Divisibilitat a l'anell de polinomis

II.1 Definició de l'anell de polinomis

Sigui K un cos commutatiu. Recordem que això vol dir que en el conjunt K hi ha definides dues operacions, que normalment anomenarem suma $(+)$ i producte $(\cdot)$, amb unes propietats que hem explicitat a (I.5). Tots els cossos que utilitzarem en aquest curs seran commutatius, és a dir, compliran $a \cdot b = b \cdot a$ per a tot parell d'elements $a, b \in K$. Per això direm simplement "cos" per indicar un cos commutatiu.

Una *successió* d'elements de K és una aplicació

$$\{0, 1, 2, \ldots\} \longrightarrow K.$$

Si indiquem per a_n la imatge de n , està clar que la successió queda determinada donant

$$(a_0, a_1, \ldots, a_n, \ldots),$$

que denotarem abreviadament per (a_n).

Un *polinomi amb coeficients a* K és una successió (a_n) amb $a_i = 0$ per a tot i llevat d'un nombre finit. Si $a_m \neq 0$ però $a_i = 0$ per a tot $i > m$ direm que m és el *grau* del polinomi (a_n): $m = \mathrm{gr}(a_n)$. Els a_i es diuen els *coeficients* del polinomi (a_n).

Observacions:

1. El polinomi $(0, 0, \ldots)$ no té cap grau.
2. Els qui recordin la noció de polinomi com una expressió del tipus $a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$, pensin que l'únic realment significatiu *a priori* en aquesta expressió són els coeficients. En la definició donada més amunt ens hem fixat simplement en ells, i hem considerat com un polinomi $(a_0, a_1, \ldots, a_n, 0, \ldots)$.

Designarem per $K[x]$ el conjunt de polinomis amb coeficients a K. El perquè d'aquesta notació quedarà clar més endavant. Definim dues operacions a $K[x]$ de la següent manera:

$$\begin{aligned}(a_0, a_1, \ldots, a_n, \ldots) + (b_0, b_1, \ldots, b_n, \ldots) = &\quad (a_0 + b_0, a_1 + b_1, \ldots, a_n + b_n, \ldots)\\ (a_0, a_1, \ldots, a_n, \ldots) \cdot (b_0, b_1, \ldots b_n, \ldots) = &\quad (a_0 b_0, a_0 b_1 + a_1 b_0, \ldots, c_n, \ldots)\end{aligned}$$

on $c_n = \sum_{i+j=n} a_i \cdot b_j$.

Amb aquestes operacions $K[x]$ és un anell commutatiu amb unitat. El zero d'aquest anell és $(0) = (0, 0, \ldots)$ i la unitat $(1, 0, \ldots)$. Es compleix també que si $(a_n) \neq (0)$,$(b_n) \neq (0)$,

$$\begin{aligned}\mathrm{gr}[(a_n) + (b_n)] \leq &\quad \mathrm{màx}[\mathrm{gr}(a_n), \mathrm{gr}(b_n)]\\ \mathrm{gr}[(a_n) \cdot (b_n)] = &\quad \mathrm{gr}(a_n) + \mathrm{gr}(b_n).\end{aligned}$$

La segona igualtat té com a conseqüències interessants:

- $(a_n) \cdot (b_n) = (0) \Rightarrow (a_n) = (0)$ o $(b_n) = (0)$.
- $(a_n) \cdot (b_n) = (a_n) \cdot (c_n) \neq 0 \Rightarrow (b_n) = (c_n)$.
- els únics elements invertibles de $K[x]$ són els de grau 0. (Demostreu-ho).

Definim ara una nova operació; si $a \in K$ i $(a_n) \in K[x]$,

$$a \cdot (a_0, a_1, a_2, \ldots) = (aa_0, aa_1, \ldots, aa_n, \ldots).$$

Això ens permet escriure

$$(a_n) = a_0 \cdot (1, 0, \ldots) + a_1 \cdot (0, 1, 0 \ldots) + a_2 \cdot (0, 0, 1, 0, \ldots) + \cdots + a_n \cdot (0, \ldots, 0, 1, 0 \ldots) + \cdots$$

Aquesta suma és sempre finita. A més a més,

$$\begin{aligned}(0, 0, 1, 0, \ldots) = &\quad (0, 1, 0, \ldots) \cdot (0, 1, 0, \ldots)\\ (0, 0, 0, 1, 0, \ldots) = &\quad (0, 0, 1, 0, \ldots) \cdot (0, 1, 0, \ldots)\\ &\quad \cdots \qquad\qquad \cdots\\ (0, 0, \overset{n}{\ldots}, 0, 1, 0, \ldots) = &\quad (0, \overset{n-1}{\ldots}, 1, 0, \ldots) \cdot (0, 1, 0, \ldots).\end{aligned}$$

Si denotem $(0, 1, 0, \ldots)$ per x obtenim

$$(a_n) = a_0 \cdot (1, 0, \ldots) + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots$$

Si a cada $a \in K$ li fem correspondre el polinomi $(a, 0, \ldots)$ obtenim una aplicació injectiva

$$K \longrightarrow K[x]$$

que conserva les operacions, és a dir:

$$\begin{aligned} a+b &\longmapsto (a+b,0,\ldots) = (a,0,\ldots)+(b,0,\ldots) \\ ab &\longmapsto (ab,0,\ldots) = (a,0,\ldots)\cdot(b,0,\ldots). \end{aligned}$$

El conjunt $\{(a,0,\ldots) \mid a \in K\}$ dels polinomis de grau 0 juntament amb el (0) està doncs en correspondència bijectiva amb K i es comporta igual que K respecte a la suma i al producte. Això justifica que designem $(a,0,\ldots)$ simplement per a. En particular $(1,0,\ldots) = 1$ i $(a_0,0,\ldots) = a_0$.

Amb aquesta notació

$$(a_n) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots$$

o abreviadament $a(x)$, que motiva la notació usual i l'expressió $K[x]$.

II.2 Divisió entera i ideals a $K[x]$

Aneu comparant aquest apartat amb el (I.1). Fixeu-vos especialment que el paper del valor absolut a $\mathbf{Z}$ aquí, a $K[x]$, el juga el grau d'un polinomi.

Teorema 2.1 (de la divisió entera) *Donats dos polinomis $a(x)$ i $b(x)$ diferents de zero de $K[x]$, existeixen dos únics polinomis $q(x)$ i $r(x)$ tals que*

$$a(x) = b(x)\cdot q(x) + r(x)$$

amb $r(x) = 0$ o $gr\, r(x) < gr\, b(x)$.

$q(x)$ i $r(x)$ es diuen el quocient *i la* resta *de la divisió entera de $a(x)$ per $b(x)$.*

DEMOSTRACIÓ: Veurem primer que existeixen $q(x)$ i $r(x)$.

Siguin

$$\begin{aligned} a(x) &= a_0 + a_1 x + \ldots + a_n x^n, \quad a_n \neq 0 \\ b(x) &= b_0 + b_1 x + \ldots + b_m x^m, \quad b_m \neq 0. \end{aligned}$$

Si $n < m$ aleshores $a(x) = b(x)\cdot 0 + a(x)$.

Si $n \geq m$ aleshores

$$a(x) - b(x)\cdot\left(\frac{a_n}{b_m}x^{n-m}\right) = r_1(x)$$

és zero o té grau $n_1 < n$.

Per conveni posem $x^0 = 1$.

Si $r_1(x) = 0$ o $n_1 < m$ tenim

$$q(x) = \frac{a_n}{b_m}x^{n-m}$$

i $r(x) = r_1(x)$.

Si $n_1 \geq m$, sigui $r_1(x) = c_0 + \ldots + c_{n_1} x^{n_1}$;

$$r_1(x) - b(x)\left(\frac{c_{n_1}}{b_m} x^{n_1 - m}\right) = r_2(x)$$

és zero o té grau $n_2 < n_1$.

Substituint més amunt obtenim

$$a(x) = b(x)\left(\frac{a_n}{b_m} x^{n-m} + \frac{c_{n_1}}{b_m} x^{n_1 - m}\right) + r_2(x).$$

Si $r_2(x) = 0$ o $n_2 < m$, aquesta expressió ens dóna un quocient i una resta. Si $n_2 \geq m$ tornem a repetir el procés restant a $r_2(x)$ un múltiple convenient de $b(x)$.

Després d'un nombre finit de passos obtindrem

$$a(x) = b(x)\left(\frac{a_n}{b_m} x^{n-m} + \frac{c_{n_1}}{b_m} x^{n_1 - m} + \ldots\right) + r_k(x)$$

amb $r_k(x) = 0$ o $\operatorname{gr} r_k(x) < m$.

Vegem ara que $q(x)$ i $r(x)$ són únics. Si

$$\begin{aligned} a(x) &= b(x) \cdot q(x) + r(x) \\ a(x) &= b(x) \cdot q_1(x) + r_1(x), \end{aligned}$$

aleshores $b(x) \cdot [q(x) - q_1(x)] = r_1(x) - r(x)$. Aquí tenim $r_1(x) - r(x) = 0$ o $\operatorname{gr}[r_1(x) - r(x)] < m$ i també $q(x) - q_1(x) = 0$ o $\operatorname{gr}\left[b(x) \cdot [q(x) - q_1(x)]\right] \geq m$.

Les segones possibilitats són incompatibles. Per tant, tindrem $r_1(x) = r(x)$ i $q_1(x) = q(x)$. □

Si la resta de la divisió entera de $a(x)$ per $b(x)$ és 0, es diu que $a(x)$ és un *múltiple* de $b(x)$ (i escriurem $a(x) = \dot{b(x)}$), o que $b(x)$ és un *divisor* de $a(x)$ (i escriurem $b(x) \mid a(x)$). Indicarem per

$$(b(x))$$

el conjunt dels múltiples de $b(x)$.

Observem que $(b(x))$ compleix les dues propietats següents:

- és tancat per la suma (és a dir, $a(x), c(x) \in (b(x)) \Rightarrow a(x) + c(x) \in (b(x))$)
- si $a(x) \in (b(x))$ i $c(x) \in K[x]$, aleshores $a(x) \cdot c(x) \in (b(x))$.

Anomenarem *ideal* de $K[x]$ tot subconjunt $I \subset K[x]$ que compleixi:

1. $a(x), b(x) \in I \Rightarrow a(x) + b(x) \in I$.
2. $a(x) \in I$ i $c(x) \in K[x] \Rightarrow a(x) \cdot c(x) \in I$.

$(b(x))$ és, doncs, un ideal. La proposició següent ens diu que tots els ideals de $K[x]$ són d'aquest tipus.

Proposició 2.2 *Si I és un ideal de $K[x]$, existeix sempre un polinomi $b(x)$ tal que $(b(x)) = I$.*

DEMOSTRACIÓ: Pot ser que $I = \{0\}$. Aleshores $I = (0)$. Si I conté algun polinomi diferent del zero, escollim-ne un de grau mínim: $b(x) \in I$. La condició 2 d'ideal ens diu que $(b(x)) \subset I$. Per altra banda, donat $a(x) \in I$ podem fer la divisió entera per $b(x)$:

$$a(x) = b(x) \cdot q(x) + r(x).$$

$r(x)$ és suma de dos polinomis de I: $r(x) = a(x) + b(x) \cdot [-q(x)]$ i per tant $r(x) \in I$.

Si $r(x) \neq 0$, tindríem un polinomi a I de grau més petit que el grau de $b(x)$, cosa que és impossible. Cal doncs que $r(x) = 0$ i $a(x) = b(x) \cdot q(x) \in (b(x))$.

Això vol dir que també $I \subset (b(x))$, i per tant $I = (b(x))$. □

Observació:

- $(b(x)) = (k \cdot b(x))$ amb $0 \neq k \in K$.
- Si $(a(x)) = (b(x))$, aleshores $a(x) = k \cdot b(x)$ amb $k \in K$. (Demostreu-ho).
- $(a(x)) \subset (b(x)) \Leftrightarrow b(x) \mid a(x)$.

II.3 Mínim comú múltiple i màxim comú divisor

La intersecció

$$(a_1(x)) \cap (a_2(x)) \cap \ldots \cap (a_n(x))$$

és el conjunt de múltiples comuns de tots els polinomis $a_1(x), \ldots, a_n(x)$. Aquest conjunt compleix les dues condicions d'ideal i per (2.2) $(a_1(x)) \cap \ldots \cap (a_n(x)) = (m(x))$ per a un cert $m(x)$. El polinomi $m(x)$ està caracteritzat per les dues propietats

- $m(x)$ és múltiple comú de $a_1(x), \ldots, a_n(x)$.
- qualsevol altre polinomi múltiple comú de $a_1(x), \ldots, a_n(x)$ és múltiple de $m(x)$.

Se l'anomena *mínim comú múltiple* de $a_1(x), \ldots, a_n(x)$:

$$m(x) = \text{m.c.m.}[a_1(x), \ldots, a_n(x)].$$

Atenció!:

Observem que $k \cdot m(x)$ (on $0 \neq k \in K$) és també mínim comú múltiple de $a_1(x), \ldots, a_n(x)$.

Considerem ara la unió: $(a_1(x)) \cup \ldots \cup (a_n(x))$. Aquest conjunt no és en general un ideal. Considerem el més petit dels ideals que contenen $(a_1(x)) \cup \ldots \cup (a_n(x))$. Aquest ha de contenir totes les sumes de múltiples de $a_1(x), \ldots, a_n(x)$. Amb això ja n'hi ha prou, ja que

$$\{a_1(x) \cdot c_1(x) + \ldots + a_n(x) \cdot c_n(x) \mid c_1(x), \ldots, c_n(x) \in K[x]\}$$

és ja un ideal de $K[x]$ que designarem per

$$(a_1(x), \ldots, a_n(x)) = \{a_1(x) \cdot c_1(x) + \ldots + a_n(x) \cdot c_n(x) \mid c_1(x), \ldots, c_n(x) \in K[x]\}.$$

Per (2.2), $(a_1(x), \ldots, a_n(x)) = (d(x))$ per a un polinomi $d(x)$ convenient. $d(x)$ està caracteritzat per

- $d(x)$ és un divisor comú de $a_1(x), \ldots, a_n(x)$, ja que

$$a_i(x) = a(x) \cdot 0 + \ldots + a_i(x) \cdot 1 + \ldots + a_n(x) \cdot 0 \in (d(x)).$$

- Tot divisor comú de $a_1(x), \ldots, a_n(x), D(x)$, divideix $d(x)$. En efecte, $D(x) \mid a_i(x)$ per a $i = 1, \ldots, n \Rightarrow (a_i(x)) \subset (D(x))$ per a $i = 1, \ldots, n \Rightarrow (d(x)) = (a_1(x), \ldots, a_n(x)) \subset (D(x))$; és a dir, $D(x) \mid d(x)$.

Direm que $d(x)$ és el *màxim comú divisor* de $a_1(x), \ldots, a_n(x)$:

$$d(x) = \text{m.c.d.}[a_1(x), \ldots, a_n(x)].$$

Observacions:

1. Si $0 \neq k \in K$, $k \cdot d(x)$ és també m.c.d.$[a_1(x), \ldots, a_n(x)]$.
2. $d(x) = a_1(x) \cdot r_1(x) + \ldots + a_n(x) \cdot r_n(x)$.

La següent proposició ens proporciona un mètode pràctic de calcular el màxim comú divisor.

Proposició 3.1 *Si $a(x) = b(x) \cdot q(x) + r(x)$ és la divisió entera de $a(x)$ per $b(x) \neq 0$, $(a(x), b(x)) = (b(x), r(x))$.*

DEMOSTRACIÓ: Compareu amb (I.2.1) i adapteu aquella demostració al cas de polinomis. □

Apliquem ara aquesta proposició fins que la resta obtinguda sigui 0:

$$\begin{aligned}
a(x) &= b(x)\cdot q(x) + r(x) && (a(x), b(x)) = (b(x), r(x)) \\
b(x) &= r(x)\cdot q_1(x) + r_1(x) && (b(x), r(x)) = (r(x), r_1(x)) \\
&\dots && \dots \\
r_{i-1}(x) &= r_i(x)\cdot q_{i+1}(x) + r_{i+1}(x) && (r_{i-1}(x), r_i(x)) = (r_i(x), r_{i+1}(x)) \\
r_i(x) &= r_{i+1}(x)\cdot q_{i+2}(x) + 0 && (r_i(x), r_{i+1}(x)) = (r_{i+1}(x), 0) = (r_{i+1}(x)).
\end{aligned}$$

Observem que $\operatorname{gr} b(x) > \operatorname{gr} r(x) > \operatorname{gr} r_1(x) > \dots$ i, per tant, sempre arriba un moment en què la resta és 0. Tenim doncs

$$(a(x), b(x)) = (r_{i+1}(x)),$$

és a dir, $r_{i+1}(x) = \text{m.c.d.}(a(x), b(x))$.

Aquest mètode de trobar el m.c.d. es coneix com *l'algorisme d'Euclides.*

Exercicis:

- $\text{m.c.d.}\big(a_1(x), \dots, a_n(x)\big) = \text{m.c.d.}\big(\text{m.c.d.}(a_1(x), \dots, a_{n-1}(x)),\ a_n(x)\big)$.
- $\text{m.c.m.}\big(a_1(x), \dots, a_n(x)\big) = \text{m.c.m.}\big(\text{m.c.m.}(a_1(x), \dots, a_{n-1}(x)),\ a_n(x)\big)$.
- Si $d(x) = \text{m.c.d.}(a(x), b(x))$ trobeu polinomis $r(x), s(x)$ tals que

$$d(x) = a(x)\cdot r(x) + b(x)\cdot s(x).$$

L'algorisme d'Euclides, juntament amb el fet que el producte del m.c.m. i el m.c.d. de dos polinomis coincideix amb el producte dels polinomis llevat de factors de K (veure apartat 4), ens dóna també una manera de calcular el m.c.m.

II.4 Polinomis irreductibles i polinomis primers entre ells

Dos polinomis $a(x)$, $b(x)$ són *primers entre ells* quan $\text{m.c.d.}(a(x), b(x)) = 1$.

Exemples:

1. $\text{m.c.d.}(x^2 - 1, x^2 + x - 6) = 1$. Observem que

$$1 = (x^2 + x - 6)\left(\frac{-1}{24}x - \frac{5}{24}\right) + (x^2 - 1)\left(\frac{1}{4} + \frac{1}{24}x\right).$$

2. Si $d(x) = \text{m.c.d.}(a(x), b(x))$ i $a(x) = d(x)\cdot r(x)$ i $b(x) = d(x)\cdot s(x)$, aleshores $\text{m.c.d.}(r(x), s(x)) = 1$.

Proposició 4.1 (Teorema d'Euclides) *Si* $a(x) \mid b(x)\cdot c(x)$ *i* $\text{m.c.d.}(a(x), b(x)) = 1$, *aleshores* $a(x) \mid c(x)$.

DEMOSTRACIÓ: Si m.c.d.$(a(x), b(x)) = 1$,

$$1 = a(x) \cdot r(x) + b(x) \cdot s(x),$$

d'on

$$c(x) = a(x) \cdot c(x) \cdot r(x) + b(x) \cdot c(x) \cdot s(x).$$

$a(x)$ divideix els dos sumands i, per tant, $a(x) \mid c(x)$. □

Proposició 4.2 *Si* $m(x) = \text{m.c.m.}(a(x), b(x))$ *i* $d(x) = \text{m.c.d.}(a(x), b(x))$, *llavors* $m(x) \cdot d(x) = k\ a(x) \cdot b(x)$ *amb* $k \in K$.

DEMOSTRACIÓ: Si $a(x) = d(x) \cdot r(x)$ i $b(x) = d(x) \cdot s(x)$ n'hi ha prou amb veure que $d(x) \cdot r(x) \cdot s(x)$ és un m.c.m.$(a(x), b(x))$. $d(x) \cdot r(x) \cdot s(x)$ és clarament múltiple de $a(x)$ i de $b(x)$. Si $M(x)$ és també un múltiple comú,

$$M(x) = a(x) \cdot c(x) = b(x) \cdot h(x);$$

és a dir, $d(x) \cdot r(x) \cdot c(x) = d(x) \cdot s(x) \cdot h(x)$.

Tenim doncs $r(x) \cdot c(x) = s(x) \cdot h(x)$, i per tant $r(x) \mid s(x) \cdot h(x)$. Com que $(r(x), s(x)) = 1$, (4.1) ens diu que $r(x) \mid h(x)$; posem $h(x) = r(x) \cdot t(x)$. Aleshores,

$$M(x) = b(x) \cdot h(x) = d(x) \cdot s(x) \cdot r(x) \cdot t(x);$$

és a dir, $M(x)$ és un múltiple de $d(x) \cdot s(x) \cdot r(x)$. □

Un polinomi $p(x)$ de grau diferent de zero es diu *irreductible* o *primer* si els seus únics divisors són k, $k \cdot p(x)$ amb $k \in K$.

Proposició 4.3 *Tot polinomi* $a(x) \neq 0$ *de grau* > 0 *és producte de polinomis irreductibles.*

DEMOSTRACIÓ: Si $a(x)$ és primer, el resultat és cert. En cas contrari, sigui $p_1(x)$ un divisor de grau mínim entre els de $a(x)$. Aleshores $p_1(x)$ és primer (ja que tots els seus divisors ho són també de $a(x)$). Posem $a(x) = p_1(x) \cdot a_1(x)$. Si $a_1(x)$ no és primer, considerem un dels seus divisors, $p_2(x)$, de grau mínim.

$p_2(x)$ és primer i

$$a(x) = p_1(x) \cdot p_2(x) \cdot a_2(x).$$

Observem que $\text{gr}\, a(x) > \text{gr}\, a_1(x) > \text{gr}\, a_2(x) > \ldots$ Arribarà un moment, doncs, en què

$$a(x) = p_1(x) \cdots p_{r-1}(x) \cdot a_r(x)$$

i $a_r(x)$ serà primer. □

Les descomposicions d'un polinomi en factors irreductibles són, fins a cert punt, úniques. Concretament:

Proposició 4.4 *Si*

$$p_1(x)\cdots p_n(x) = q_1(x)\cdots q_m(x)$$

i tots els factors són polinomis irreductibles, llavors $n = m$ *i els polinomis* $\{p_i(x)\}$ *són els mateixos que els* $\{q_j(x)\}$ *llevat de factors del cos* K.

DEMOSTRACIÓ: Procedim per inducció sobre n. Si $n = 1$, clarament $m = 1$ i $p_1(x) = q_1(x)$.

Suposem ara que el resultat és cert sempre que $n \leq r-1$. Donada l'expressió

$$p_1(x)\cdots p_r(x) = q_1(x)\cdots q_m(x),$$

tenim que $p_r(x)$ divideix

$$q_1(x)(q_2(x)\cdots q_m(x)).$$

Si $p_r(x)$ no coincideix amb $q_1(x)$ (llevat de factors de K), aleshores és primer amb ell i, per tant, $p_r(x)$ divideix $q_2(x)\cdots q_m(x)$.

Repetint el raonament arribarem a un $q_j(x)$ que serà igual a $p_r(x)$ llevat d'un factor de K; $p_r(x) = q_j(x)\cdot k$, $0 \neq k \in K$. Suprimint aquest factor comú tenim

$$p_1(x)\cdots p_{r-1}(x) = q_1(x)\cdots q_{j-1}(x)\cdot(k^{-1}q_{j+1}(x))\cdots q_m(x)$$

i podem aplicar la hipòtesi d'inducció per deduir que també

$$p_1(x),\ldots,p_{r-1}(x)$$

coincideixen amb els $q_i(x)$ restants (llevat de factors de K) i, en particular, que $r-1 = m-1$, d'on $r = m$. □

II.5 Zeros d'un polinomi

Si $a(x) = a_0 + a_1x + \ldots + a_nx^n$ és un polinomi de $K[x]$ i $k \in K$, anomenarem *valor de* $a(x)$ *a* k

$$a(k) = a_0 + a_1k + \ldots + a_nk^n \in K.$$

Observem que el valor de $a(x) + b(x)$ a k és $a(k) + b(k)$, i el valor de $a(x)\cdot b(x)$ a k és $a(k)\cdot b(k)$.

Si $a(k) = 0$ direm que k és un *zero* o una *arrel* de $a(x)$.

Proposició 5.1 *k és un zero del polinomi $a(x) \neq 0$ si i només si $a(x)$ és divisible per $x-k$.*

DEMOSTRACIÓ: Si $\operatorname{gr} a(x) = 0$, tindrem $a(x) = a_0 \neq 0$ i k no serà un zero : $a(k) = a_0 \neq 0$.

Si $\operatorname{gr} a(x) \geq 1$, fem la divisió entera per $(x-k)$. La resta haurà d'ésser 0 o de grau 0:

$$a(x) = (x-k)\cdot q(x) + r \quad r \in K.$$

Aleshores, $0 = a(k) = r$ i per tant $x - k$ divideix $a(x)$.□

Direm que $k \in K$ és un *zero de multiplicitat* p del polinomi $a(x) \in K[x]$ si $a(x) = (x-k)^p b(x)$ i $b(k) \neq 0$; és a dir, si $a(x)$ és divisible per $(x-k)^p$ però no ho és per $(x-k)^{p+1}$.

Corol·lari 5.2 *Si* $gr\, a(x) = n$, *la suma de les multiplicitats dels zeros de* $a(x)$ *és* $\leq n$.□

Pot ser que dos polinomis diferents $a(x) \neq b(x)$ prenguin el mateix valor sobre tots els $k \in K$? Tindríem aleshores un polinomi $a(x) - b(x) \neq 0$ tots els $k \in K$ del qual serien zeros. (5.2) ens diu que si K té prou elements ($> \mathrm{gr}(a(x) - b(x))$) això no és possible. En particular:

Proposició 5.3 *Si* K *és infinit i* $a(k) = b(k)$ *per a tot* $k \in K$, *llavors* $a(x) = b(x)$.□

Exemple:

Considerem els polinomis $a(x) = x-2$ i $b(x) = x^3-2$ amb coeficients a $\mathbf{Z}/(3)$. Com a polinomis, $a(x) \neq b(x)$; ara bé, $a(0) = -2 = b(0)$, $a(1) = -1 = b(1)$, $a(2) = 0 = b(2)$.

Acabarem aquest apartat donant un criteri molt senzill per a trobar els zeros racionals d'un polinomi de $\mathbf{Q}[x]$. Considerem

$$a(x) = a_0 + a_1 x + \ldots + a_n x^n$$

amb

$$a_0, \ldots, a_n \in \mathbf{Q}.$$

Podem sempre trobar un enter $m \neq 0$ tal que $ma_0, \ldots, ma_n \in \mathbf{Z}$. El polinomi $ma(x)$ té els mateixos zeros que $a(x)$ i els coeficients enters. El problema queda reduït, doncs, a trobar els zeros racionals d'un polinomi amb coeficients enters.

Considerem, doncs,

$$b(x) = b_0 + b_1 x + \ldots + b_n x^n$$

amb

$$b_0, \ldots, b_n \in \mathbf{Z}.$$

Sigui p/q un zero de $b(x)$ amb p,q primers entre ells. De

$$b_0 + b_1 \frac{p}{q} + \ldots + b_n \frac{p^n}{q^n} = 0$$

obtenim

$$b_0 q^n + b_1 p q^{n-1} + \ldots + b_{n-1} p^{n-1} q + b_n p^n = 0.$$

Com que m.c.d.$(p,q) = 1$, aplicant el teorema d'Euclides resulta $p \mid b_0$ i $q \mid b_n$. Per tant,

Proposició 5.4 *Si p/q, amb p, q primers entre ells, és un zero del polinomi*

$$b(x) = b_0 + b_1x + \ldots + b_nx^n \in \mathbf{Z}[x],$$

aleshores $p \mid b_0$ i $q \mid b_n$.□

Corol·lari 5.5 *Si $k \in \mathbf{Z}$ és un zero de*

$$b_0 + b_1x + \ldots b_nx^n \in \mathbf{Z}[x],$$

aleshores $k \mid b_0$.□

II.6 Polinomis irreductibles de $\mathbf{R}[x]$

L'estudi dels polinomis irreductibles a $\mathbf{R}[x]$ i a $\mathbf{C}[x]$ es basa en el següent teorema:

Teorema 6.1 (fonamental de l'Àlgebra) *Tot polinomi de grau ≥ 1 amb coeficients complexos té un zero.*

No donarem la demostració d'aquest teorema, que va més enllà de l'objectiu del llibre. Però sí que en farem alguns comentaris i en traurem conseqüències. Primer de tot diguem que, malgrat el seu nom, no es tracta d'un teorema "algebraic" sinó d'un teorema "topològic"; en altres paraules, aquest teorema és conseqüència de les propietats de completitud de $\mathbf{C}$ (i de $\mathbf{R}$) i no de les propietats de les seves operacions. Observem també que del teorema es dedueix que tot polinomi de $\mathbf{C}[x]$ de grau ≥ 1 és producte de factors lineals (de grau 1) i, per tant,

Corol·lari 6.2 *Els polinomis irreductibles de $\mathbf{C}[x]$ són els de grau 1.*

Aquest corol·lari dóna, de fet, un altre enunciat del teorema, ja que si $\alpha + \beta x$ és un factor lineal del polinomi $a(x)$, $a(x)$ té el zero $-\alpha/\beta$.

Estudiarem ara els polinomis irreductibles de $\mathbf{R}[x]$. Tot polinomi real $a(x) = a_0 + a_1x + \ldots + a_nx^n$ pot considerar-se també com un polinomi amb coeficients complexos. En general, si $a(x) = a_0 + a_1x + \ldots + a_nx^n$ és de $\mathbf{C}[x]$ indicarem per $\bar{a}(x)$

$$\bar{a}_0 + \bar{a}_1x + \ldots + \bar{a}_nx^n.$$

Si $z = a + bi \in \mathbf{C}$, $\bar{z} = a - bi$ indica el seu conjugat. Aleshores $a(x)$ té coeficients reals si i només si

$$\bar{a}(x) = a(x).$$

Per altra banda, si $z \in \mathbf{C}$

$$\bar{a}(\bar{z}) = \bar{a}_0 + \bar{a}_1\bar{z} + \ldots + \bar{a}_n\bar{z}^n = \overline{(a_0 + a_1z + \ldots + a_nz^n)} = \overline{a(z)}$$

i, en particular, si z és un zero de $a(x)$, $(a(z) = 0)$, $\bar{z}$ és zero de $\bar{a}(x)$, $(\bar{a}(\bar{z}) = 0)$.

Quan $a(x)$ té coeficients reals, resulta que sempre que z sigui un zero, $\bar{z}$ també ho és. Aleshores, o bé $z = \bar{z}$, és a dir, z és un zero real de $a(x)$, o bé $z \neq \bar{z}$ i $a(x)$ és divisible per

$$(x - z)(x - \bar{z}) = x^2 - (z + \bar{z})x + z\bar{z},$$

que és un polinomi amb coeficients reals. A més a més, $x^2 - (z + \bar{z})x + z\bar{z}$ és irreductible a $\mathbf{R}[x]$, ja que en cas contrari tindria un divisor de primer grau i per tant un zero real.

Així doncs els polinomis irreductibles de $\mathbf{R}[x]$ són de grau ≤ 2.

Nota:

A l'anell $\mathbf{Q}[x]$ es poden trobar polinomis irreductibles de grau tan gran com es vulgui.

II.7 Els anells $K[x]/(m(x))$

Sigui $m(x)$ un polinomi de $K[x]$. Direm que dos polinomis $a(x)$ i $b(x)$ són *congruents mòdul* $m(x)$ si $a(x) - b(x) \in (m(x))$. Això equival a dir que les restes de les divisions enteres de $a(x)$ i $b(x)$ per $m(x)$ són iguals (compareu amb I.4). Escriurem aleshores

$$a(x) \equiv b(x) \quad (m(x)).$$

Aquesta relació és clarament d'equivalència. Designem per $[a(x)]$ la classe d'equivalència de $a(x)$, és a dir, el conjunt de polinomis congruents amb $a(x)$ mòdul $m(x)$.

El conjunt d'aquestes classes d'equivalència el denotarem per

$$K[x]/(m(x))$$

i en direm *quocient* de $K[x]$ per $(m(x))$. Observem que hi ha tantes classes d'equivalència com restes possibles en les divisions enteres per $m(x)$. Aquestes restes són precisament els polinomis de grau més petit que el grau de $m(x)$. En altres paraules, a cada classe d'equivalència hi ha un i només un polinomi de grau més petit que el de $m(x)$.

En el conjunt $K[x]/(m(x))$ podem definir dues operacions, suma:

$$[a(x)] + [b(x)] = [a(x) + b(x)],$$

i producte:

$$[a(x)] \cdot [b(x)] = [a(x) \cdot b(x)].$$

Cal veure però que la classe suma i la classe producte no depenen dels representants $a(x)$, $b(x)$ escollits. És a dir, que si $a'(x)$, $b'(x)$ són uns altres representants de $[a(x)], [b(x)]$ respectivament, llavors

$$[a'(x) + b'(x)] = [a(x) + b(x)] \qquad [a'(x) \cdot b'(x)] = [a(x) \cdot b(x)].$$

La comprovació es fa exactament igual que en el cas de les classes de restes a $\mathbf{Z}$ (I.5).

$K[x]/(m(x))$ té, amb aquestes operacions, l'estructura d'un anell commutatiu amb unitat; ara bé, aquest anell posseeix algunes propietats que no tenia $K[x]$. Per exemple,

Proposició 7.1 *Si* $(a(x), m(x)) = (1)$, $[a(x)]$ *té un invers a* $K[x]/(m(x))$.
Si $(a(x), m(x)) = (d(x))$ *amb* $gr\, d(x) \geq 1$, $[a(x)]$ *és un divisor de 0 a* $K[x]/(m(x))$.

La demostració és anàloga a la de I.5.1.□

En particular:

Corol·lari 7.2 *Si* $p(x) \in K[x]$ *és irreductible,* $K[x]/(p(x))$ *és un cos.*□

Si $\mathrm{gr}\, p(x) \geq 1$,

$$\begin{array}{ccc} K & \longrightarrow & K[x]/(p(x)) \\ k & \mapsto & [k] \end{array}$$

és injectiva i conserva la suma i el producte. Aquest fet justifica que denotem els elements $[k]$ simplement per k i el subconjunt de $K[x]/(p(x))$ imatge de l'aplicació per la lletra K. Amb aquesta notació escriurem

$$K \subset K[x]/(p(x)).$$

Quan $p(x)$ és irreductible obtenim, doncs, un cos $K[x]/(p(x))$ que "conté" K.

Tot polinomi $a(x)$ amb coeficients a K pot considerar-se també com un polinomi amb coeficients a $K[x]/(p(x))$. En particular, el polinomi

$$p(x) = p_0 + p_1 x + \ldots + p_n x^n$$

pot considerar-se amb coeficients a $K[x]/(p(x))$.

Tenim, aleshores, que si posem $\alpha = [x] \in K[x]/(p(x))$

$$p(\alpha) = p([x]) = p_0 + p_1[x] + \ldots + p_n[x^n] = [p_0 + p_1 x + \ldots + p_n x^n] = [0];$$

és a dir, el polinomi $p(x)$, que era irreductible a $K[x]$, té un zero (i, per tant, té un divisor lineal) a $K[x]/(p(x))$.

El cos $K[x]/(p(x))$ es denota per $K(\alpha)$ i es diu una *extensió algebraica* de K.

Exemple:

Considerem $x^2 + 1$, que és un polinomi irreductible a $\mathbf{R}[x]$. Els elements del cos $\mathbf{R}[x]/(x^2+1)$ tenen, cada un, un i només un representant de primer grau

$$[a + bx].$$

La suma de dues classes és

$$[a + bx] + [c + dx] = [(a + c) + (b + d)x]$$

i el producte

$$[a+bx]\cdot[c+dx]=[ac+(ad+bc)x+bdx^2]=[(ac-bd)+(ad+bc)x].$$

Aleshores, si "identifiquem" cada $a\in\mathbf{R}$ amb $[a]\in\mathbf{R}[x]/(x^2+1)$ i denotem $[x]$ per i, obtenim que els elements del quocient són de la forma

$$[a+bx]=a+b[x]=a+bi,$$

i amb aquesta notació

$$i^2=[x^2]=[-1]=-1$$

i les dues operacions s'expressen així:

$$\begin{aligned}(a+bi)+(c+di)&=(a+c)+(b+d)i\\(a+bi)\cdot(c+di)&=(ac-bd)+(ad+bc)i.\end{aligned}$$

Existeix, doncs, una correspondència bijectiva entre el cos $\mathbf{R}[x]/(x^2+1)$ i el cos $\mathbf{C}$ dels nombres complexos, que conserva les operacions. Podem dir, doncs, que el cos $\mathbf{R}[x]/(x^2+1)$ no és altra cosa que el cos $\mathbf{C}$ dels nombres complexos. Amb més precisió, es diu que $\mathbf{R}[x]/(x^2+1)$ i $\mathbf{C}$ són dos cossos *isomorfs*.

Exemple:

Considerem

$$\mathbf{Q}[x]/(x^2-2)=K;$$

x^2-2 és irreductible a $\mathbf{Q}[x]$ i, per tant, K és un cos que conté $\mathbf{Q}$. Tot element de $\mathbf{Q}$ té un representant (i només un) de primer grau $ax+b$.

Les dues operacions són

$$\begin{aligned}[ax+b]+[cx+d]&=[(a+c)x+(b+d)]\\[ax+b]\cdot[cx+d]&=[acx^2+(ad+bc)x+bd]=[(ad+bc)x+2ac+bd].\end{aligned}$$

L'element $\alpha=[x]\in K$ compleix $\alpha^2=2$ i és un zero del polinomi $X^2-2\in K[X]$. L'extensió algebraica $K=\mathbf{Q}(\alpha)$ és isomorfa a la seva imatge per l'aplicació

$$\begin{array}{ccc}\mathbf{Q}(\alpha)&\longrightarrow&\mathbf{R}\\a\alpha+b&\mapsto&a\sqrt{2}+b.\end{array}$$

II.8 Nota històrica

El simbolisme emprat en els polinomis i equacions s'ha anat elaborant al llarg de la història i no assolí la forma actual fins a començaments del segle 18. Sembla ésser que els signes "+" i "−" foren usats per primera vegada per J. Widman el segle 16 desplaçant les lletres "p" i "m", abreviacions de "plus" i "minus". François Viète (1540–1603), un parlamentari que dedicava el temps lliure a les matemàtiques, va donar un gran impuls a l'àlgebra simbòlica, utilitzant lletres (les primeres de l'abecedari) per a les variables. La nostra equació "$5BA^2 - 2CA + A^3 = D$" l'escrivia "B5 in A quadratum − C plano 2 in A + A cubum aequator D solido" (i això fou un gran avenç sobre els seus predecessors). L'obra de René Descartes (1596–1650) conté ja la notació actual amb dues variants menors: "xx" per "x^2" i "$\propto$" per "=". En la resolució d'equacions, i especialment en el que fa referència als apartats 5 i 6, cal esmentar Paolo Ruffini (1765–1822) i Carl Friedrich Gauss (1777–1855), el "príncep de les matemàtiques" segons la inscripció feta gravar pel rei George V de Hannover, que demostrà el teorema fonamental de l'algebra i en donà quatre demostracions tot buscant-ne una purament algebraica.

Augustin-Louis Cauchy (1789–1857) fou el primer a observar, el 1847, que els nombres complexos es poden considerar com classes d'equivalència de $\mathbf{R}[x]$ mòdul $x^2 + 1$. És curios que, malgrat que des de Gauss ja es treballa amb relacions d'equivalència a $\mathbf{Z}$ i a $K[x]$, passa gairebé tot el segle 19 fins que s'introdueix sistemàticament el conjunt quocient corresponent.

Dos matemàtics destaquen per les seves aportacions inicials a la teoria de cossos, considerant extensions d'un cos per una arrel d'un polinomi: Niels Henrik Abel (1802–1829) i Évariste Galois (1811–1831), ambdós estudiant la resolubilitat de les equacions de grau ≥ 5. Tant Abel com Galois moriren molt joves i tràgicament; l'un tuberculós i en la misèria, l'altre en un duel.

II.9 Exercicis

1. Calculeu el màxim comú divisor $d(x)$ dels polinomis $p(x) = x^5 - 5x^3 + 4x$ i $q(x) = x^3 - 2x^2 - 5x + 6$. Trobeu dos polinomis $a(x)$ i $b(x)$ de manera que $p(x) \cdot a(x) + q(x) \cdot b(x) = d(x)$.

2. Si $p, q \in \mathbf{R}[x]$ són tals que $(p, q) = (1)$, demostreu que $(p + q, p \cdot q) = (1)$.

3. Factoritzeu com a producte de polinomis irreductibles:

   a) $x^3 - 2$, $x^{12} - 4$, $x^p - 1$ a $\mathbf{Q}[x]$, $\mathbf{R}[x]$ i $\mathbf{C}[x]$.

   b) $x^p - x$ a $\mathbf{Z}/(p)\,[x]$, $x^3 + 2x^2 + 5x + 1$ a $\mathbf{Z}/(7)\,[x]$.

4. Determineu un polinomi $p(x)$ de grau mínim tal que $x^2 + 1 \mid p(x)$ i $x^3 + 1 \mid p(x) - 1$.

5. Si $p(x) \in \mathbf{Z}[x]$ i $p(r/s) = 0$ amb $(r, s) = 1$, demostreu que $r - s \mid p(1)$ i $r + s \mid p(-1)$.

6. Calculeu tots els zeros racionals de $20x^3 - 56x^2 - 33x + 9$ i de $12x^5 - 17x^4 + 7x^3 - 5x^2 - 22x - 5$.

7. Determineu un polinomi $p(x)$ de grau 5 tal que $p(0) = p(1) = p(2) = p(3) = p(4) = 1$.

8. Descomponeu $x^4 + a^2 \in \mathbf{R}[x]$ en factors irreductibles.

9. Descomponeu $(x+1)^n + (x-1)^n \in \mathbf{C}[x]$ en factors lineals.

10. Demostreu que $2 + \sqrt[3]{3}$, $\sqrt{2} + \sqrt{3}$, $\sqrt[3]{2} + \sqrt{3}$ i $\sqrt{2} + \sqrt{3} + \sqrt{5}$ són cada un d'ells zero d'un polinomi de $\mathbf{Z}[x]$. Determineu aquests polinomis.

11. Racionalitzeu les expressions

$$\frac{1}{\sqrt[4]{8} - \sqrt[4]{4} + \sqrt[4]{2} - 3} \quad \text{i} \quad \frac{1}{\sqrt[3]{7} + \sqrt[4]{7}}.$$

12. Donat el polinomi $p(x) = 3x^3 + 5x^2 + 5x + 2$,

a) Trobeu tots els zeros complexos de $p(x)$,

b) Determineu els divisors de zero dels anells $\mathbf{C}[x]/(p(x))$ i $\mathbf{R}[x]/(p(x))$.

13. Resoleu a $\mathbf{R}[x]/(x^2 + 2x + 1)$ l'equació $z^2 + z + 1/4 = 0$.

14. Per a quins valors $a \in \mathbf{C}$ l'equació $z^2 + z + a = 0$ a l'anell $A = \mathbf{C}[x]/(x^3)$ té infinites solucions?

15. Per a cada element $a \in A = \mathbf{R}[x]/(x^2 + x)$ determineu quantes solucions té l'equació $z^2 = a$. Representeu l'anell A sobre el pla i dividiu-lo en regions segons el nombre de solucions de l'equació anterior.

II.10 Exercicis de programar

16. Resolució a $\mathbf{Z}/(p)$ de l'equació de segon grau

$$ax^2 + bx + c = 0.$$

(Indicació: utilitzeu com a subrutines accessòries els exercicis I.20 i I.21).

17. Divisió entera de dos polinomis de $\mathbf{Z}/(p)[x]$. (Indicació: utilitzeu com a subrutina accessòria l'exercici I.20).

18. Factorització d'un polinomi de $\mathbf{Z}/(p)[x]$ en polinomis irreductibles. (Indicació: utilitzeu com a subrutines accessòries els exercicis I.20, I.21 dins de II.16 i II.17).

19. Càlcul dels zeros racionals d'un polinomi de $\mathbf{Q}[x]$. (Indicació: utilitzeu la proposició II.5.4 i l'exercici II.5).

Capítol III

Grups

III.1 Definició i exemples

Un *grup* és un conjunt G juntament amb una operació $\cdot$ que compleix les propietats:

- Associativa: $g \cdot (g' \cdot g'') = (g \cdot g') \cdot g'' \qquad \forall g, g', g'' \in G.$
- Existeix un element e, que anomenarem *element neutre*, tal que

$$g \cdot e = g = e \cdot g \qquad \forall g \in G.$$

- Per a cada $g \in G$ hi ha un element, que anomenarem l'*invers de* g i denotarem per g^{-1}, tal que

$$g \cdot g^{-1} = e = g^{-1} \cdot g.$$

Si es compleix també la propietat commutativa:

$$g \cdot g' = g' \cdot g \qquad \forall g, g' \in G,$$

direm que el *grup* és *commutatiu* o *abelià*. En aquest cas l'operació es denota sovint per $+$, l'element neutre per 0 (i es diu *zero*) i l'element invers per $-g$ (i es diu l'*oposat* de g).

Quan indiquem l'operació per $\cdot$, notació multiplicativa, l'element neutre s'acostuma a dir *unitat* i a escriure 1. Amb aquesta notació multiplicativa és costum suprimir el punt que indica l'operació i escriure simplement gg' per indicar $g \cdot g'$.

Exemples:

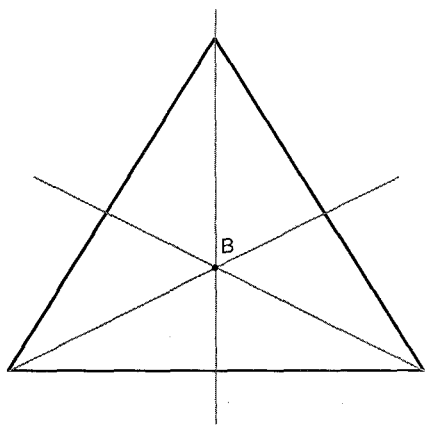
1. Els nombres enters $\mathbf{Z}$ amb la suma formen un grup commutatiu. El mateix val per als racionals $\mathbf{Q}$ i els reals $\mathbf{R}$. Els nombres naturals $\mathbf{N} = \{1, 2, \ldots\}$ no són grup amb la suma.

Els nombres racionals no nuls, $\mathbf{Q} - \{0\}$, amb el producte formen un grup commutatiu. El mateix val per a $\mathbf{R} - \{0\}$. Ni $\mathbf{Z} - \{0\}$ ni $\mathbf{N}$ no són grup amb el producte.

2. Els nombres complexos $\mathbf{C}$ amb la suma són un grup commutatiu. $\mathbf{C} - \{0\}$ amb el producte és un grup commutatiu.

   $\mathcal{S}^1 = \{z \in \mathbf{C} \mid |z| = 1\}$ amb el producte, és un grup commutatiu.

3. Tots els exemples anteriors són grups commutatius. Els exemples més senzills de grups *no* commutatius sorgeixen a la geometria en estudiar determinats conjunts de moviments. Així, per exemple, el conjunt de moviments del pla que deixen fix un triangle equilàter està format per tres simetries respecte a eixos que passen per un vèrtex i el punt mig del costat oposat, els girs de 120° i 240° al voltant del baricentre, i la identitat o gir de 0°. En aquests exemples



geomètrics l'operació és *la composició*: la composició dels moviments g' i g és el moviment $g \circ g'$ que resulta de fer successivament els moviments g' i g. (Atenció a l'ordre!). Aquesta operació no és commutativa.

Aquests grups de moviments sortiran de manera natural en estudiar la geometria. Ara, a l'apartat 2, ens anem a ocupar d'uns altres grups no commutatius senzills: els grups de permutacions.

III.2 Permutacions

Sigui $A = \{a_1, \ldots, a_n\}$ un conjunt amb n elements. Una *permutació de* A és una aplicació bijectiva

$$\sigma : A \longrightarrow A.$$

La composició de permutacions és, clarament, una permutació. A més a més es compleixen les propietats:

- Associativa: $\sigma \circ (\rho \circ \tau) = (\sigma \circ \rho) \circ \tau \qquad \forall \sigma, \rho, \tau.$
- Hi ha una permutació I tal que: $\sigma \circ I = \sigma = I \circ \sigma \qquad \forall \sigma.$
  I és l'aplicació identitat: $I(a) = a \qquad \forall a \in A.$
- Per a tota permutació σ hi ha una permutació σ^{-1} tal que $\sigma \circ \sigma^{-1} = I = \sigma^{-1} \circ \sigma$. σ^{-1} és l'aplicació inversa de σ.

El conjunt $\mathcal{S}_A$ de les permutacions de A amb la composició és, doncs, un grup. De la composició en direm també *producte* i escriurem, a vegades, $\sigma\tau$ per $\sigma \circ \tau$.

Per simplificar la notació considerarem d'ara endavant, si no diem el contrari, que $A = \{1, \ldots, n\}$. El conjunt de permutacions de $\{1, \ldots, n\}$ el designarem per $\mathcal{S}_n$. Si A és qualsevol conjunt amb n elements, les propietats de $\mathcal{S}_A$ són exactament les mateixes que les de $\mathcal{S}_n$ (veure apartat 4).

Per donar una permutació concreta σ, haurem de dir quines són les imatges de cada un dels elements $1, 2, \ldots, n$. Una manera còmoda de fer-ho és escrivint aquests elements en fila i davall de cada un la seva imatge:

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Proposició 2.1 *Si $n \geq 3$, $\mathcal{S}_n$ no és commutatiu.*

DEMOSTRACIÓ: En efecte, es té

$$\begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 3 & 2 & 1 & 4 \ldots n \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 1 & 3 & 2 & 4 \ldots n \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 3 & 1 & 2 & 4 \ldots n \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 1 & 3 & 2 & 4 \ldots n \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 3 & 2 & 1 & 4 \ldots n \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 2 & 3 & 1 & 4 \ldots n \end{pmatrix}. \square$$

Per $n = 2$, $\mathcal{S}_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}$ és un grup commutatiu.

Un element j es diu *fix* per una permutació σ si $\sigma(j) = j$. Si j no és fix, formem la successió

$$j, \sigma(j), \sigma^2(j), \ldots, \sigma^r(j), \ldots$$

on $\sigma^2 = \sigma \circ \sigma$ i, en general, $\sigma^r = \sigma \circ \sigma^{r-1}$. Com que $\{1, 2, \ldots, n\}$ és finit, en algun moment un element $\sigma^k(j)$ coincidirà amb un dels anteriors. El primer element que torna a sortir és precisament el j: en efecte, si $\sigma^k(j) = \sigma^h(j)$ amb $h < k$, per ser σ bijectiva,

$$\sigma^k(j) = \sigma^h(j) \Rightarrow \sigma^{k-1}(j) = \sigma^{h-1}(j) \Rightarrow \ldots \Rightarrow \sigma^{k-h}(j) = j\ ;$$

és a dir, j hauria sortit ja.

Siguin, doncs,

$$j, \sigma(j), \ldots, \sigma^{r-1}(j)$$

diferents i $\sigma^r(j) = j$. Direm que la permutació σ és un *cicle d'ordre* r si deixa fixos tots els elements que no apareixen en la successió anterior. Escriurem aleshores

$$\sigma = (j, \sigma(j), \sigma^2(j), \ldots, \sigma^{r-1}(j)).$$

Exemple:

A $\mathcal{S}_5$,

$$(2,3,1,5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 4 & 2 \end{pmatrix}.$$

En el cas que la permutació σ no sigui un cicle, considerem un j_1 no fix per σ i diferent de $j, \sigma(j), \ldots, \sigma^{r-1}(j)$. Siguin

$$j_1, \sigma(j_1), \ldots, \sigma^{r_1-1}(j_1)$$

diferents i $\sigma^{r_1}(j_1) = j_1$. Un moment de reflexió ens convencerà que cap d'aquests elements no havia sortit en la successió $j, \sigma(j), \ldots$. Repetim aquest procés tantes vegades com sigui necessari fins a esgotar tots els elements no fixos per σ. A cada pas agafem un element j_m no fix i que no hagi sortit abans i formem la successió

$$j_m, \sigma(j_m), \ldots, \sigma^{r_m-1}(j_m); \quad \sigma^{r_m}(j_m) = j_m.$$

Els seus elements són tots diferents dels que hem obtingut amb anterioritat. Suposem que, després de fer aquest pas, no queda ja cap altre element no fix. Aleshores σ és producte de cicles

$$(j_m, \sigma(j_m), \ldots, \sigma^{r_m-1}(j_m)) \ldots (j_1, \sigma(j_1), \ldots, \sigma^{r_1-1}(j_1))(j, \sigma(j), \ldots, \sigma^{r-1}(j)).$$

Observem que aquests cicles commuten entre ells, ja que afecten elements diferents. D'aquesta forma queda demostrada la proposició següent:

Proposició 2.2 *Tota permutació és producte de cicles.*□

Els cicles d'ordre 2 es diuen *transposicions.*

Proposició 2.3 *Tot cicle és producte de transposicions.*

DEMOSTRACIÓ:

$$(a_1, a_2, \ldots, a_m) = (a_1, a_2)(a_2, a_3) \ldots (a_{m-1}, a_m).\square$$

Exemples:

1. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 7 & 5 & 2 & 6 & 3 \end{pmatrix} = (1,4,5,2)(3,7) = (1,4)(4,5)(5,2)(3,7)$

2. $I = (1,2)(1,2) = (3,4)(2,3)(1,2)(2,4)(3,2)(1,4).$

La identitat I, i per tant qualsevol permutació, es pot posar de moltes maneres com a producte de transposicions. Anem a veure, però, que el nombre de transposicions en aquests productes té sempre la mateixa paritat.

Proposició 2.4 *La permutació identitat no es pot posar com a producte d'un nombre senar de transposicions.*

DEMOSTRACIÓ: La demostració que anem a fer es basa en un fet aparentment anecdòtic: si en l'expressió del producte

$$P = \prod_{i,j}(j-i)$$

on $1 \leq i < j \leq n, \quad i,j \in \{1,2,\ldots,n\}$, permutem les i,j segons una transposició, obtenim la mateixa expressió amb signe canviat. Ho explicarem. Si σ és una permutació de $\{1,2,\ldots,n\}$, escriurem

$$\sigma P = \prod_{i,j}(\sigma(j) - \sigma(i)).$$

En el cas que $\sigma = (h,k)$, $h < k$, quins són els factors de σP ? Observem que

- si i,j són diferents de h,k, $\quad \sigma(j) - \sigma(i) = j - i$;
- si $i < h < k$, el factor $h - i$ de P passa a ésser $k - i$ a σP, el factor $k - i$ de P passa a ésser $h - i$ a σP; és a dir, l'únic canvi en aquest cas és un canvi en la posició dels factors;
- si $h < k < j$, el factor $j - h$ de P passa a ésser $j - k$ a σP, el factor $j - k$ de P passa a ésser $j - h$ a σP; com abans, només hi ha hagut un canvi de posició;
- si $h < i < k$, el factor $i - h$ de P passa a ésser $i - k$ a σP, el factor $k - i$ de P passa a ésser $h - i$ a σP; ara, el canvi és de posició i de signe; el signe canvia, però, dues vegades i, per tant, no afecta el producte. Finalment,
- si $i = h < k = j$, el factor $k - h$ de P passa a ésser $h - k$ a σP. Aquest és l'únic canvi de signe que afecta el producte.

Obtenim, doncs,

$$\sigma P = -P.$$

Suposem ara que

$$I = \tau_n \circ \ldots \circ \tau_2 \circ \tau_1,$$

on les τ_i són transposicions. Apliquem a P successivament les transposicions $\tau_1, \tau_2, \ldots, \tau_n$. Obtindrem $(-1)^n P$. Per altra banda, aplicar $\tau_1, \ldots, \tau_n$ equival a aplicar la identitat i, per tant, el resultat ha d'ésser P. És a dir, $(-1)^n P = P$, d'on resulta que n és parell. □

Corol·lari 2.5 *Si $\sigma = \tau_p \circ \ldots \circ \tau_1 = \rho_q \circ \ldots \circ \rho_1$ són dues descomposicions de la permutació σ com a producte de transposicions, aleshores p i q són de la mateixa paritat.*

DEMOSTRACIÓ: Multiplicant els dos productes per ρ_1 a la dreta i tenint en compte que $\rho_1 \circ \rho_1 = I$, obtenim

$$\tau_p \circ \ldots \circ \tau_1 \circ \rho_1 = \rho_q \circ \ldots \circ \rho_2.$$

Multipliquem a la dreta per $\rho_2, \ldots, \rho_q$ successivament; obtenim

$$\tau_p \circ \ldots \circ \tau_1 \circ \rho_1 \circ \ldots \circ \rho_q = I.$$

(2.4) ens diu aleshores que $p + q$ és parell, i, per tant, p i q són tots dos parells o tots dos senars. □

Una permutació es diu *parella* si descompon en un nombre parell de transposicions; una permutació es diu *senar* si descompon en un nombre senar.

El producte de transposicions parelles i senars segueix la regla dels signes: el producte de dues permutacions parelles o de dues senars, és parell; el producte d'una permutació parella i una senar, és senar. Aquest fet motiva l'assignació a les permutacions parelles del signe "+" i a les permutacions senars del signe "−". Anomenarem *aplicació signe* l'aplicació

$$\varepsilon : \mathcal{S}_n \longrightarrow \{+1, -1\}$$

tal que

$$\varepsilon(\sigma) = +1 \quad \text{si } \sigma \text{ és parell,}$$
$$\varepsilon(\sigma) = -1 \quad \text{si } \sigma \text{ és senar.}$$

Es compleix

$$\varepsilon(I) = 1, \quad \varepsilon(\sigma \circ \tau) = \varepsilon(\sigma) \cdot \varepsilon(\tau).$$

En particular, per (2.3), si $(a_1, \ldots, a_m)$ és un cicle d'ordre m,

$$\varepsilon(a_1, \ldots, a_m) = (-1)^{m-1}.$$

III.3 Subgrups

Sigui S un subconjunt no buit d'un grup G. Si es compleix que
(1) per a tot parell $g, g' \in S$, $gg' \in S$,
l'operació de G dóna lloc a una operació a S, que anomenarem "l'operació induïda" per la de G. Ens interessen els subconjunts S de G que compleixen (1) i que amb l'operació induïda siguin un grup al seu torn. D'aquests subconjunts en direm "subgrups".

Suposem, doncs, que S compleix (1) i té, per tant, una operació induïda. Aquesta operació serà automàticament associativa (per ser-ho la de G); si té un element neutre $e' \in S$,

$$\forall g \in S \quad e'g = g;$$

multiplicant a la dreta per l'invers de g (a G), obtenim $e' = e$. És a dir, si S té element neutre, aquest ha d'ésser el mateix element neutre e de G. De manera semblant es veu que si un element g de S té invers per l'operació induïda a S, aquest ha de coincidir amb l'invers g^{-1} que g té a G. Per tant, les condicions que ha de complir S per a ésser un grup són:

- $e \in S$.
- $g \in S \Rightarrow g^{-1} \in S$.

La primera d'aquestes dues condicions és conseqüència de la segona i de (1). Hem justificat així la definició següent de subgrup:
Un subconjunt S, no buit, d'un grup G direm que és un *subgrup de* G, si compleix

1. $g, g' \in S \Rightarrow gg' \in S$,
2. $g \in S \Rightarrow g^{-1} \in S$.

De fet aquestes dues condicions es poden sintetitzar en una:

Proposició 3.1 *Un subconjunt $S \neq \emptyset$ d'un grup G és un subgrup de G si i només si compleix*

$$g', g \in S \Rightarrow g'g^{-1} \in S.$$

DEMOSTRACIÓ: Si S és subgrup,

$$g', g \in S \Rightarrow g', g^{-1} \in S \Rightarrow g'g^{-1} \in S.$$

Si S compleix la condició de l'enunciat i $g \in S$ és qualsevol, $e = gg^{-1} \in S$. Aleshores, $\forall g \in S$, $g^{-1} = eg^{-1} \in S$. Això demostra 2, que, ara, ens permet demostrar 1: $g', g \in S \Rightarrow g', g^{-1} \in S \Rightarrow g'g \in S$, ja que $(g^{-1})^{-1} = g$.□

Exemples:

1. $S^1 = \{z \in \mathbf{C} \mid |z| = 1\}$ és un subgrup de $\mathbf{C} - \{0\}$ amb el producte.

2. $\mathbf{Z}$ amb $+$ és un grup. Si S és un subgrup de $\mathbf{Z}$,la primera condició de la definició de subgrup ens diu que si $a, b \in S$, aleshores $a + b \in S$. Si $a \in S$ i $n \in \mathbf{Z}$, aleshores na és 0 o suma d'unes quantes a o suma d'unes quantes $-a$. En tots els casos, $na \in S$. Així doncs, S és un ideal de $\mathbf{Z}$ i, per (I.1.2), és de la forma $S = (m)$.

3. Estudiem els subgrups del grup de permutacions $\mathcal{S}_3$. Els elements de $\mathcal{S}_3$ són

$$I,\ A = (1,2,3),\ B = (1,3,2),\ \tau_1 = (2,3),\ \tau_2 = (1,3),\ \tau_3 = (1,2).$$

Tot subgrup ha de contenir I. Per tant, amb un element hi ha només un subgrup: $\{I\}$. Els conjunts

$$\{I, \tau_1\},\ \{I, \tau_2\} \text{ i } \{I, \tau_3\}$$

són subgrups. En canvi, ni $\{I, A\}$ ni $\{I, B\}$ no ho són; de fet, si un subgrup conté A, ha de contenir $A^2 = B$ i si un subgrup conté B ha de contenir $B^2 = A$. El conjunt

$$\{I, A, B\}$$

és un subgrup. Cap altre subconjunt propi de $\mathcal{S}_3$ no és subgrup. En altres paraules, si un subgrup S de $\mathcal{S}_3$ conté a més de I dues permutacions que no siguin A i B, aleshores $S = \mathcal{S}_3$. Considerem per exemple el cas en què $\tau_1 \in S$ i $A \in S$; aleshores

$$B = A^2 \in S, \quad \tau_2 = \tau_1 A \in S, \quad \tau_3 = A\tau_1 \in S$$

i S conté tots els elements. De manera semblant es veu la resta de casos.

Donat un subconjunt S d'un grup G, anomenarem *subgrup generat per* S el "més petit" subgrup de G que conté S. El designarem per (S). Aquí, "més petit" vol dir que (S) està contingut en qualsevol altre subgrup que contingui S. Cal preguntar-se, però, si existeix sempre (S) i, en tal cas, com es forma.

Observem, primer, que (S) ha de contenir els elements de S, els inversos dels elements de S i els productes d'uns i altres. No cal afegir més elements; el conjunt de productes

$$\{s_1 \dots s_n \mid s_i \in S \text{ o } s_i^{-1} \in S\}$$

és ja un subgrup. (Demostreu-ho!). Aquest conjunt és, doncs, (S).

Exemples:

1. El conjunt $(a_1, \dots, a_n)$ construït a (I.2) és precisament el subgrup generat per $\{a_1, \dots, a_n\}$.

2. El subgrup de $\mathcal{S}_3$ generat per $\{I, \tau_1, A\}$ és tot $\mathcal{S}_3$.

III.4 Homomorfismes

Siguin G i G' dos grups. Una aplicació

$$f : G \longrightarrow G'$$

es diu un *homomorfisme* (o *morfisme*) *de grups* si, $\forall g_1, g_2 \in G$, es compleix que $f(g_1 \cdot g_2) = f(g_1) \cdot f(g_2)$.

Exemples:

1. Considerem el grup dels nombres reals amb la suma, $(\mathbf{R}, +)$, i el grup dels nombres reals positius amb el producte, $(\mathbf{R}^+, \cdot)$. L'aplicació

$$\begin{array}{ccc} \mathbf{R} & \longrightarrow & \mathbf{R}^+ \\ x & \longmapsto & e^x \end{array}$$

és un homomorfisme de grups, ja que $e^{x+y} = e^x \cdot e^y$.

2. L'aplicació signe definida a l'apartat 2

$$\varepsilon : S_n \longrightarrow \{+1, -1\}$$

és un homomorfisme.

Proposició 4.1 *Sigui $f : G \longrightarrow G'$ un homomorfisme de grups. Siguin e i e' els elements neutres de G i G' respectivament. Aleshores,*

a) $f(e) = e'$,

b) $f(g^{-1}) = (f(g))^{-1}, \quad \forall g \in G$.

DEMOSTRACIÓ: a) Si $g \in G$, $f(g) = f(ge) = f(g)f(e)$, d'on $f(e) = e'$.
b) $f(g)f(g^{-1}) = f(gg^{-1}) = f(e) = e'$, d'on $f(g^{-1}) = (f(g))^{-1}$. □

Proposició 4.2 *Si $f : G \longrightarrow G'$ i $h : G' \longrightarrow G''$ són dos homomorfismes de grups, aleshores $h \circ f : G \longrightarrow G''$ és també un homomorfisme.*

Demostreu-ho. □

Un homomorfisme injectiu es diu un *monomorfisme*; un morfisme exhaustiu es diu un *epimorfisme*; un morfisme bijectiu es diu un *isomorfisme*; si $f : G \longrightarrow G'$ és un isomorfisme, direm que G i G' són *isomorfs* i escriurem $G \cong G'$.

Dos grups isomorfs tenen les mateixes propietats, "els mateixos subgrups", etc.

Exercici:

Si A i B són dos conjunts de n elements, demostreu que $\mathcal{S}_A$ i $\mathcal{S}_B$ són isomorfs.

Anomenarem *nucli d'un homomorfisme* $f : G \longrightarrow G'$

$$\operatorname{Nuc} f = \{g \in G \mid f(g) = e'\}.$$

Anomenarem *imatge d'un homomorfisme* f

$$\operatorname{Im} f = \{g' \in G' \mid \text{ existeix } g \in G,\ f(g) = g'\}.$$

És fàcil comprovar que $\operatorname{Nuc} f$ és un subgrup de G i $\operatorname{Im} f$ és un subgrup de G'.

Proposició 4.3 *Sigui $f : G \longrightarrow G'$ un homomorfisme de grups.*

a) f és injectiva si i només si $\operatorname{Nuc} f = \{e\}$.

b) f és exhaustiva si i només si $\operatorname{Im} f = G'$.

DEMOSTRACIÓ: La segona afirmació no és altra cosa que la definició d'exhaustivitat. Demostrem, doncs, la primera.

Si $g \in \operatorname{Nuc} f$, aleshores $f(g) = e' = f(e)$ (per (4.1)); per ser f injectiva, això implica que $g = e$.

Recíprocament, suposem que $f(g_1) = f(g_2)$; aleshores

$$f(g_1 g_2^{-1}) = f(g_1) f(g_2)^{-1} = e',$$

d'on $g_1 g_2^{-1} \in \operatorname{Nuc} f = \{e\}$, i, per tant, $g_1 g_2^{-1} = e$; és a dir, $g_1 = g_2$.□

III.5 Grup quocient. Subgrups normals

Recordem que els quocients $\mathbf{Z}/(m)$ (I.4) estaven definits a partir de la relació d'equivalència: $a \equiv b \Leftrightarrow a - b \in (m)$. En general, si G és un grup commutatiu amb una operació $+$ i H és un subgrup de G, la relació "$g_1 \sim g_2 \Leftrightarrow g_1 - g_2 \in H$" és d'equivalència. Quan G no és commutatiu hi ha dues possibles generalitzacions:

I. $g_1 \sim g_2 \Leftrightarrow g_1 \cdot g_2^{-1} \in H$

II. $g_1 \approx g_2 \Leftrightarrow g_2^{-1} \cdot g_1 \in H$.

Tant una com l'altra són relacions d'equivalència. Demostrem-ho per a la primera:

- és reflexiva: $g \in G, gg^{-1} = e \in H$, d'on $g \sim g$,
- és simètrica: $g_1 \sim g_2 \Rightarrow g_1 g_2^{-1} \in H \Rightarrow g_2 g_1^{-1} = (g_1 g_2^{-1})^{-1} \in H \Rightarrow g_2 \sim g_1$,
- és transitiva: $g_1 \sim g_2,\ g_2 \sim g_3 \Rightarrow g_1 g_2^{-1} \in H,\ g_2 g_3^{-1} \in H \Rightarrow g_1 g_3^{-1} = (g_1 g_2^{-1})(g_2 g_3^{-1}) \in H \Rightarrow g_1 \sim g_3$.

Les classes d'equivalència per a aquesta relació són

$$[g] = \{g_1 \in G \mid g_1 \sim g\} = \{g_1 \in G \mid g_1 = hg,\ h \in H\}.$$

Posarem $[g] = Hg$ i el conjunt quocient el denotarem per $H\backslash G$.

Les classes d'equivalència per a la relació II són

$$\{g\} = \{g_1 \in G \mid g_1 \approx g\} = \{g_1 \in G \mid g_1 = gh,\ h \in H\}.$$

Posarem $\{g\} = gH$ i el conjunt quocient el denotarem per G/H.

Una manera lògica de definir les operacions a $H\backslash G$ i a G/H seria

$$[g_1][g_2] = [g_1 g_2] \text{ i } \{g_1\}\{g_2\} = \{g_1 g_2\},$$

respectivament. Aquesta no és sempre, però, una bona definició; vegem-ho en un exemple.

Exemple:

Considerem el subgrup $H = \{I, \tau_1\}$ de $G = \mathcal{S}_3$. Amb les notacions de l'apartat 3, les classes de $H\backslash G$ són

$$H = \{I, \tau_1\},\ HA = \{A, \tau_2\},\ HB = \{B, \tau_3\}.$$

Les classes de G/H són

$$H = \{I, \tau_1\},\ AH = \{A, \tau_3\},\ BH = \{B, \tau_2\}.$$

El producte de les classes HA i HB s'hauria d'obtenir fent el producte d'un representant de HA per un de HB. Ara bé,

$$AB = I, \text{ d'on el producte seria } H = [I]$$
$$\tau_2\tau_3 = A, \text{ d'on el producte seria } HA = [A].$$

L'operació no queda, doncs, ben determinada. Una cosa semblant passa amb les classes de G/H.

En quines condicions el producte $[g_1][g_2] = [g_1 g_2]$ està ben determinat? Sempre que per a tot parell $h_1, h_2 \in H$

$$(h_1 g_1)(h_2 g_2) = h g_1 g_2$$

per a un cert $h \in H$. Ara bé,

$$h_1 g_1 h_2 g_2 = h g_1 g_2 \Leftrightarrow h_1 g_1 h_2 g_1^{-1} = h \Leftrightarrow g_1 h_2 g_1^{-1} = h_1^{-1} h \in H.$$

Per tant, existeix un tal $h \in H$ si i només si $g_1 h_2 g_1^{-1} \in H$. L'operació està, doncs, ben definida si per a tot $g_1 \in G$ i $h_2 \in H$, $g_1 h_2 g_1^{-1} \in H$; és a dir, si

$$g_1 H g_1^{-1} \subset H \quad \forall g_1 \in G,$$

on $g_1 H g_1^{-1} = \{g_1 h_2 g_1^{-1} \mid h_2 \in H\}$. En particular, per a tot $g \in G$, aplicant la inclusió anterior a g i a g^{-1}, tenim

$$gHg^{-1} \subset H$$
$$g^{-1}Hg \subset H, \text{ és a dir } H \subset gHg^{-1},$$

d'on resulta

$$gHg^{-1} = H.$$

Direm que H és un *subgrup normal* si és un subgrup que compleix la igualtat anterior per a tot $g \in G$. Si H és un subgrup normal de G, en el conjunt $H\backslash G$ hi ha una operació ben definida: $[g_1][g_2] = [g_1 g_2]$. Amb aquesta operació $H\backslash G$ és un grup.

Observem que $gHg^{-1} = H$ equival a

$$gH = Hg \quad \forall g \in G;$$

és a dir, si H és normal, les classes per a les dues relacions I i II coincideixen, i, per tant, els conjunts quocients també, $H\backslash G = G/H$. Naturalment, en aquest cas, l'operació de G/H també està ben definida i coincideix amb la de $H\backslash G$.

Nota:

El mateix resultat s'obté, naturalment, si comencem estudiant en quines condicions l'operació de G/H està ben definida.

Exercici:

Demostreu que si H és normal, G/H és un grup. El seu element neutre és $e = H$; l'invers de $[g]$ és $[g^{-1}]$.

Sigui H un subgrup normal de G. L'aplicació

$$\begin{array}{ccc} G & \longrightarrow & G/H \\ g & \longmapsto & [g] \end{array}$$

és un epimorfisme de nucli H.

Proposició 5.1 *Un subgrup H és normal si i només si és nucli d'un homomorfisme.*

DEMOSTRACIÓ: Tot subgrup normal H d'un grup G és el nucli de l'epimorfisme $G \longrightarrow G/H$ que acabem de definir.

Suposem, ara, que $H = \operatorname{Nuc} f$, per a un cert homomorfisme de grups $f : G \longrightarrow G'$. Per a tot $g \in G$ i tot $h \in H$

$$f(ghg^{-1}) = f(g)f(h)f(g^{-1}) = f(g)f(g)^{-1} = e,$$

d'on resulta que $ghg^{-1} \in \operatorname{Nuc} f = H$ i, per tant, H és un subgrup normal. □

Exemple:

El conjunt $\mathcal{A}_n$ de les permutacions parelles de $\mathcal{S}_n$ és un subgrup normal, ja que és el nucli de l'aplicació signe

$$\varepsilon : \mathcal{S}_n \longrightarrow \{+1, -1\}.$$

$\mathcal{A}_n = \operatorname{Nuc} \varepsilon$ es diu el *grup alternat d'ordre n*. En particular, $\mathcal{A}_3 = \{I, A, B\}$ és un subgrup normal de $\mathcal{S}_3$.

Teorema 5.2 (d'isomorfisme) *Sigui $f : G \longrightarrow G'$ un homomorfisme de grups; aleshores*

$$G/\operatorname{Nuc} f \cong \operatorname{Im} f.$$

DEMOSTRACIÓ: Per (5.1), $\operatorname{Nuc} f$ és normal i $G/\operatorname{Nuc} f$ és un grup. Tots els elements d'una mateixa classe de $G/\operatorname{Nuc} f$ tenen la mateixa imatge per f; en efecte, si $gh \in [g]$ amb $h \in \operatorname{Nuc} f$, aleshores

$$f(gh) = f(g)f(h) = f(g).$$

Podem definir, doncs, una aplicació

$$\begin{array}{ccc} G/\operatorname{Nuc} f & \longrightarrow & \operatorname{Im} f \\ g & \longmapsto & f(g). \end{array}$$

Aquesta aplicació és, clarament, un homomorfisme exhaustiu. Per veure que és injectiu, només cal comprovar que l'única classe que s'aplica a l'element neutre és $[e] = \operatorname{Nuc} f$ (4.3); en efecte, si $f(g) = e'$, $g \in \operatorname{Nuc} f$ i, per tant, $[g] = \operatorname{Nuc} f$.□

III.6 Producte directe de grups

Es diu que un grup G és *producte directe* dels seus subgrups H_1 i H_2 si

a) H_1 i H_2 són subgrups normals de G;

b) $H_1 \cap H_2 = \{e\}$ (on e és l'element neutre de G);

c) $G = \{h_1 h_2 \mid h_1 \in H_1,\ h_2 \in H_2\}$.

Proposició 6.1 *G és producte directe dels seus subgrups H_1 i H_2 si i només si*

1. *tot $g \in G$ s'expressa de manera única com a producte $g = h_1 h_2$ amb $h_1 \in H_1$ i $h_2 \in H_2$;*
2. *$h_1 h_2 = h_2 h_1$, $\forall h_1 \in H_1$, $\forall h_2 \in H_2$.*

DEMOSTRACIÓ: Suposem que G és producte directe de H_1 i H_2. Aleshores es compleix 1, ja que si $g = h_1h_2 = h_1'h_2'$ amb $h_1, h_1' \in H_1$ i $h_2, h_2' \in H_2$, $h_1^{-1}h_1' = h_2(h_2')^{-1} \in H_1 \cap H_2 = \{e\}$, d'on resulta que $h_1^{-1}h_1' = e$ i $h_2(h_2')^{-1} = e$, i, per tant, $h_1 = h_1'$ i $h_2 = h_2'$.

També es compleix 2, ja que

$$(h_1h_2)(h_2h_1)^{-1} = (h_1h_2)(h_1^{-1}h_2^{-1});$$

però, per ésser H_1 normal, $h_2h_1^{-1}h_2^{-1} \in H_1$, d'on $h_1(h_2h_1^{-1}h_2^{-1}) \in H_1$ i, per ésser H_2 normal, $h_1h_2h_1^{-1} \in H_2$, d'on $(h_1h_2h_1^{-1})h_2^{-1} \in H_2$. Ara bé, com que $H_1 \cap H_2 = \{e\}$, cal que $(h_1h_2)(h_2h_1)^{-1} = h_1h_2h_1^{-1}h_2^{-1} = e$; és a dir, $h_1h_2 = h_2h_1$.

Recíprocament, suposem, ara, que G compleix 1 i 2. La condició c) és part de 1 i, per tant, ja es compleix. Per demostrar que H_1 és normal, considerem elements qualssevol $g = h_1h_2 \in G$ ($h_1 \in H_1$, $h_2 \in H_2$) i $h_1' \in H_1$; fent servir 2 obtenim

$$gh_1'g^{-1} = h_1h_2h_1'h_2^{-1}h_1^{-1} = h_1h_2h_2^{-1}h_1'h_1^{-1} = h_1h_1'h_1^{-1} \in H_1.$$

Anàlogament es veu que H_2 és normal.

Finalment, per provar b), suposem que $h \in H_1 \cap H_2$; per 1, les dues expressions $he = eh$ han d'ésser la mateixa; així doncs, $h = e$. □

Siguin ara G_1 i G_2 dos grups (poden ésser el mateix). Anomenarem *producte directe de G_1 i G_2* el conjunt $G_1 \times G_2$ juntament amb l'operació

$$(g_1, g_2)(g_1', g_2') = (g_1g_1', g_2g_2') \quad \forall g_1, g_1' \in G_1,\ g_2, g_2' \in G_2.$$

Com quasi sempre que parlem en abstracte, fem servir la notació multiplicativa per a G_1 i G_2. S'entén, però, que en cada cas particular els elements g_1 i g_1' s'operen amb l'operació concreta de G_1, i els elements g_2 i g_2' amb l'operació concreta de G_2.

El fet que s'utilitzi el nom de "producte directe" també per al grup $G_1 \times G_2$ que acabem de definir no és casual. La següent proposició ho justifica.

Proposició 6.2 *Sigui $G_1 \times G_2$ el producte directe dels grups G_1 i G_2. Existeixen dos subgrups de $G_1 \times G_2$, G_1' i G_2', isomorfs a G_1 i G_2 respectivament, tals que $G_1 \times G_2$ és el producte directe de G_1' i G_2'.*

DEMOSTRACIÓ: Agafem

$$G_1' = \{(g_1, e) \in G_1 \times G_2\} \qquad G_2' = \{(e, g_2) \in G_1 \times G_2\},$$

on indiquem per e tant l'element neutre de G_1 com el de G_2. És fàcil veure que G_1' i G_2' són subgrups de $G_1 \times G_2$ i que les aplicacions

$$\begin{array}{ccc} G_1 & \longrightarrow & G_1' \\ g_1 & \longmapsto & (g_1, e) \end{array} \qquad \begin{array}{ccc} G_2 & \longrightarrow & G_2' \\ g_2 & \longmapsto & (e, g_2) \end{array}$$

són isomorfismes. Comprovem que $G_1 \times G_2$ és producte directe de G_1' i G_2' veient que compleix les condicions 1 i 2 de (6.1):
Tot element $(g_1, g_2) \in G_1 \times G_2$ es pot posar com

$$(g_1, g_2) = (g_1, e)(e, g_2)$$

de manera única; això demostra 1. A més, es té sempre

$$(g_1, e)(e, g_2) = (e, g_2)(g_1, e),$$

que demostra 2.□

Exemple:

Estudiem el producte directe $\mathbf{Z}/(a) \times \mathbf{Z}/(b)$, quan a i b són primers entre ells (I.5). (A $\mathbf{Z}/(a)$ i a $\mathbf{Z}/(b)$ considerem l'operació suma). Sumant l'element $([1],[1])$ amb ell mateix prou vegades podem obtenir tots els elements de $\mathbf{Z}/(a) \times \mathbf{Z}/(b)$: en efecte, això equival a dir que tot element $([m],[q])$ és de la forma

$$([m],[q]) = ([1],[1]) + \overset{n}{\cdots} + ([1],[1]) = ([n],[n]);$$

n ha d'ésser, doncs, tal que

$$n = m + at = q + br.$$

Com que a i b són primers entre ells, existeixen t i r tals que $m - q = -at + br$ i, per tant, existeix el nombre n que buscàvem.

Aquest fet ens porta de manera natural a definir l'aplicació exhaustiva

$$\begin{array}{ccc} \mathbf{Z} & \longrightarrow & \mathbf{Z}/(a) \times \mathbf{Z}/(b) \\ n & \longmapsto & ([n],[n]). \end{array}$$

Aquesta aplicació és, clarament, un morfisme de nucli $(a) \cap (b)$; però, com que m.c.d.$(a,b) = 1$, el seu mínim comú múltiple és ab : $(a) \cap (b) = (ab)$. El teorema 5.2 ens diu aleshores que

$$\mathbf{Z}/(ab) \cong \mathbf{Z}/(a) \times \mathbf{Z}/(b).$$

III.7 Grups cíclics

Un grup G es diu *cíclic* si està generat per un element g (que es diu un *generador* de G). Escriurem

$$G = (g).$$

Tal com hem vist a l'apartat 3, el subgrup generat per g està format per g, g^{-1} i productes d'aquests elements: g^n, $(g^{-1})^n$. Fent servir la notació $g^0 = e$, $g^{-n} = (g^{-1})^n$, tenim, doncs,

$$(g) = \{g^n \mid n \in \mathbf{Z}\}.$$

Exemples:

1. $\mathbf{Z}$ amb la suma és un grup cíclic generat per 1.

2. $\mathbf{Z}/(m)$ amb la suma és un grup cíclic generat per $[1]$.

La següent proposició ens diu que aquests són els únics exemples de grups cíclics, llevat d'isomorfismes.

Proposició 7.1 *Tot grup cíclic és isomorf a* $\mathbf{Z}$ *o a un* $\mathbf{Z}/(m)$.

DEMOSTRACIÓ: Sigui $G = (g) = \{g^n \mid n \in \mathbf{Z}\}$ un grup cíclic. L'operació a G és $g^n g^m = g^{n+m}$, que ens diu que l'aplicació

$$\begin{array}{ccc} \mathbf{Z} & \longrightarrow & G \\ n & \longmapsto & g^n \end{array}$$

és un epimorfisme. El seu nucli és un subgrup de $\mathbf{Z}$ que, com hem vist a l'exemple 2 de l'apartat 3, serà de la forma (m). Aleshores (5.2) ens diu que $\mathbf{Z}/(m) \cong G$. El cas $m = 0$ correspon a $\mathbf{Z} \cong G$. □

Un grup cíclic $G = (g)$ es diu d'*ordre* $m \neq 0$ si és isomorf a $\mathbf{Z}/(m)$; en aquest cas, $g^n = e$ sempre que $n = \dot{m}$. Un grup cíclic $G = (g)$ es diu d'*ordre infinit* si és isomorf a $\mathbf{Z}$; llavors $g^n = e$ només quan $n = 0$.

Proposició 7.2 *Tot subgrup d'un grup cíclic és cíclic.*

DEMOSTRACIÓ: Sigui S un subgrup de $G = (g)$.

Si $S = \{e\} = (e)$, S és cíclic.

Si $S \neq \{e\}$, sigui g^k un element de S. Aleshores, també, $g^{-k} \in S$ i, per tant, S conté potències de g amb exponent positiu. Sigui $g^m \in S$ amb exponent positiu mínim. S conté el subgrup generat per g^m : $(g^m) \subset S$. Anem a veure que $(g^m) = S$; en efecte, sigui $g^k \in S$ i fem la divisió entera de k per m:

$$k = mq + r \text{ amb } 0 \leq r < m.$$

Aleshores

$$g^r = g^{k-mq} = g^k (g^m)^{-q} \in S,$$

ja que $g^k \in S$ i $(g^m)^{-q} \in (g^m) \subset S$. Com que l'exponent m era mínim, cal que $r = 0$ i, per tant, $k = \dot{m}$, d'on resulta que $g^k \in (g^m)$.□

III.8 Grups finits

Anomenarem *ordre d'un grup finit* G el nombre dels seus elements, i el denotarem per $|G|$. Observem que si G és cíclic aquest ordre coincideix amb el definit a l'apartat anterior.

Anomenarem *ordre d'un element* $g \in G$ l'ordre del subgrup cíclic generat per g.

Proposició 8.1 *Si S és un subgrup del grup finit G, $|S|$ divideix $|G|$.*

DEMOSTRACIÓ: Formem el conjunt quocient G/S (apartat 5). Totes les classes gS tenen el mateix nombre d'elements que S, ja que l'aplicació $S \to gS$ tal que $x \longmapsto gx$ és bijectiva; per tant, si a G/S hi ha i classes, $|G| = i \cdot |S|$.□

De la demostració de (8.1) es dedueix que el nombre de classes de G/S i de $S\backslash G$ és el mateix; en direm *índex de S a G* i el denotarem per $[G : S]$. Per (8.1),

$$[G : S] = |G|/|S|.$$

Corol·lari 8.2 *L'ordre d'un element divideix l'ordre del grup.*□

Corol·lari 8.3 *Si $|G| = p$ és primer, G és un grup cíclic d'ordre p.*

DEMOSTRACIÓ: Sigui $g \in G$, $g \neq e$. Per (8.2) g és d'ordre 1 o p. Si g fos d'ordre 1, g seria igual a e; així doncs, g és d'ordre p. Per tant, $|(g)| = p = |G|$, d'on resulta $(g) = G$.□

Exemple:

Recordem que en estudiar els subgrups de $\mathcal{S}_3$ a l'apartat 3, hem trobat subgrups d'ordre $1, 2, 3$ i 6, que són els divisors de $|\mathcal{S}_3| = 6$. Respecte a l'ordre dels elements, I és d'ordre 1, τ_1, τ_2 i τ_3 són d'ordre 2 i A i B són d'ordre 3. No hi ha cap element d'ordre 6.

Un dels objectius de la teoria de grups finits és determinar quins grups hi ha de cada ordre (llevat d'isomorfismes, naturalment). El problema no està ni de bon tros resolt, encara que existeixen llistes de tots els grups finits fins a ordres molt elevats (per exemple, en el llibre *Group Tables* de A.D. Thomas i G.V. Wood). El corol·lari 8.3 dóna una informació important: d'un ordre primer p hi ha un únic grup i amb una estructura molt simple: $\mathbf{Z}/(p)$.

Estudiem, ara, quants grups hi ha d'ordre 4.

D'entrada podem considerar dos casos:

I. Hi ha un element d'ordre 4. Aleshores es tracta del grup cíclic $\mathbf{Z}/(4)$.

II. Tots els elements són d'ordre 2 (llevat de l'element neutre, que sempre és d'ordre 1). Sigui, doncs, $G = \{e, a, b, c\}$ amb les relacions $a^2 = b^2 = c^2 = e$. Formem una taula o quadre on escriurem tots els productes de dos elements de G: a cada lloc escriurem el producte de l'element situat a la mateixa "altura" a la primera columna i el situat a la mateixa "altura" a la primera fila (en aquest ordre). Aquest quadre es diu la "taula de l'operació" de G. En aquest cas II tenim, de moment,

	e	a	b	c
e	e	a	b	c
a	a	e		
b	b		e	
c	c			e.

Els elements que surten en una mateixa fila (o en una mateixa columna) han d'ésser tots diferents (per què ?!). Per tant, el producte ab ha d'ésser forçosament c. Pel mateix motiu resulta $ac = b$, $ba = c, \ldots$ Obtenim així com a única taula possible

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e.

Observem que aquest grup també és commutatiu. A més, es veu fàcilment que és el producte directe dels seus subgrups $\{e, a\}$ i $\{e, b\}$. Aquests subgrups són isomorfs a $\mathbf{Z}/(2)$. L'aplicació

$$G \longrightarrow \mathbf{Z}/(2) \times \mathbf{Z}/(2)$$

que aplica e, a, b, c en $([0],[0]), ([1],[0]), ([0],[1]), ([1],[1])$, respectivament, és un isomorfisme de grups.

Hem demostrat així, la

Proposició 8.4 *Només hi ha dos grups d'ordre* 4 : $\mathbf{Z}/(4)$ *i* $\mathbf{Z}/(2)\times\mathbf{Z}/(2)$. *Ambdós són commutatius.*□

Passem a estudiar, ara, els grups d'ordre 6. Ens podem trobar amb els tres casos següents (que no s'exclouen):

I. Hi ha un element d'ordre 6. Aleshores es tracta del grup cíclic $\mathbf{Z}/(6)$.

II. Hi ha un element g d'ordre 3. Per (8.1), $[G : (g)] = 2$. En general, tenim

Proposició 8.5 *Tot subgrup S d'un grup G d'índex 2 és normal.*

DEMOSTRACIÓ: Si S té índex 2, G/S consta de dos elements S i $gS = G - S$. Anàlogament, $S \setminus G$ consta de dos elements , S i $Sg = G - S$. Per tant, les classes per la dreta i per l'esquerra són les mateixes i S és un subgrup normal.□

Tornem al cas II. (g) és normal i $G/(g)$ és un grup amb 2 elements:

$$[e] = (g) = \{e, g, g^2\} \qquad \text{i} \qquad [g_1] = g_1(g) = \{g_1, g_1g, g_1g^2\}.$$

L'element $[g_1]$ ha d'ésser d'ordre 2; és a dir, $g_1^2 \in [e] = \{e, g, g^2\}$.

- Si $g_1^2 = g$, aleshores $g_1^3 = gg_1 \neq e$, ja que en cas contrari $g_1 = g^{-1} = g^2$ i $g_1 \notin (g)$; així doncs, l'ordre de g_1 és més gran que 3, i per tant, és 6 (per (8.2)). G és, doncs, cíclic i isomorf a $\mathbf{Z}/(6)$.
- Si $g_1^2 = g^2$, aleshores $g_1^3 = g^2g_1 \neq e$, ja que en cas contrari $g_1 = g^{-2} = g$ i $g_1 \notin (g)$. Així doncs, l'ordre de g_1 és més gran que 3 i, com abans, $G \cong \mathbf{Z}/(6)$.

Si $g_1^2 = e$, aleshores, com que $\{g_1, g_1g, g_1g^2\} = \{g_1, gg_1, g^2g_1\}$ (ja que $g_1(g) = (g)g_1$), tenim

- o bé $g_1g = gg_1$ (i $g_1g^2 = g^2g_1$), d'on $(g_1g)^2 = g_1^2g^2 = g^2 \neq e$ i $(g_1g)^3 = g_1^3g^3 = g_1 \neq e$; és a dir, l'ordre de g_1g és més gran que 3 i, per tant, com abans, $G \cong \mathbf{Z}/(6)$,
- o bé $g_1g = g^2g_1$ (i $g_1g^2 = gg_1$). En aquest cas, la taula d'operació del grup és

	e	g	g^2	g_1	g_1g	g_1g^2
e	e	g	g^2	g_1	g_1g	g_1g^2
g	g	g^2	e	g_1g^2	g_1	g_1g
g^2	g^2	e	g	g_1g	g_1g^2	g_1
g_1	g_1	g_1g	g_1g^2	e	g	g^2
g_1g	g_1g	g_1g^2	g_1	g^2	e	g
g_1g^2	g_1g^2	g_1	g_1g	g	g^2	e.

Observem que aquesta taula és "la mateixa" que la del grup de permutacions $\mathcal{S}_3$, canviant només g, g^2 per A, B i g_1, g_1g, g_1g^2 per les tres transposicions τ_1, τ_2, τ_3. En altres paraules, l'aplicació

$$G \longrightarrow \mathcal{S}_3$$

que porta e, g, g^2, g_1, g_1g, g_1g^2 a I, A, B, τ_1, τ_2, τ_3, respectivament, és un isomorfisme: $G \cong \mathcal{S}_3$.

III. L'últim cas a considerar és quan tots els elements de G són d'ordre 2.

Proposició 8.6 *Si tots els elements d'un grup G són d'ordre 2, G és un grup commutatiu.*

DEMOSTRACIÓ: Un element és d'ordre 2 si i només si és invers d'ell mateix. Per tant, per a tot parell $g_1, g_2 \in G$, $(g_1g_2)(g_2g_1)^{-1} = g_1g_2g_2g_1 = g_1g_1 = e$, d'on $g_1g_2 = g_2g_1$.□

Considerem, doncs, $g \in G$; (g) és normal per ésser G commutatiu. Formem el grup $G/(g)$, que tindrà 3 elements. Per (8.3), $G/(g)$ és cíclic. Sigui $[g_1]$ un generador de $G/(g)$; $[g_1]$ hauria d'ésser d'ordre 3, però $[g_1]^2 = [g_1^2] = [e]$. Aquesta contradicció ens diu que aquest tercer cas no es pot presentar mai.

Hem demostrat, doncs,

Proposició 8.7 *Només hi ha dos grups d'ordre 6: un de commutatiu, $\mathbf{Z}/(6)$, i un de no commutatiu, $\mathcal{S}_3$.*□

Acabem donant sense demostració un resultat molt important. Hem dit abans que el problema de determinar tots els grups d'un cert ordre no estava resolt. Nogensmenys, sí que ho està el problema de determinar tots els grups commutatius d'un cert ordre.

Encara més, es coneixen tots els grups commutatius amb un nombre finit de generadors. Concretament tenim

Teorema 8.8 (d'estructura dels grups commutatius) *Tot grup commutatiu G amb un nombre finit de generadors és producte directe d'un nombre finit de grups $\mathbf{Z}$ i $\mathbf{Z}/(m)$:*

$$G \cong \mathbf{Z} \times \ldots \times \mathbf{Z} \times \mathbf{Z}/(m_1) \times \ldots \times \mathbf{Z}/(m_r).$$

Aquesta descomposició és única si $m_1 \mid m_2 \mid \ldots \mid m_r$.

Existeixen demostracions elementals d'aquest teorema. (Veure, per exemple, el llibre *Algebra*, volum I, segona edició, de P. M. Cohn (John Wiley & Sons, 1982)).

III.9 Nota històrica

Els inicis de la teoria de grups es poden situar en l'estudi que Joseph-Louis Lagrange (1736–1813) féu de la resolució de les equacions de grau n. La idea de Lagrange fou escollir una funció racional de les arrels de l'equació que fos invariant per a totes les permutacions de les arrels. Encara que el mètode de Lagrange no dóna el resultat esperat, sí que motiva l'aparició dels primers resultats de la teoria de grups de permutacions (i, fent abstracció, de grups finits).

Carl Friedrich Gauss (1777–1855), a les seves *Disquisitiones Arithmeticae*, dedica una secció a l'estudi de les formes quadràtiques $ax^2 + 2bxy + cy^2$; defineix una composició de formes, una relació d'equivalència entre elles i comprova que les classes d'equivalència tenen l'estructura d'un grup commutatiu (evidentment, no utilitza aquest llenguatge!). És raonable pensar que Gauss tenia ja la idea del teorema d'estructura dels grups abelians finits, encara que no fou demostrat fins l'any 1870 per Leopold Kronecker (1823–1891).

És, però, el treball d'Évariste Galois (1811–1832) que dóna un impuls extraordinari a la teoria de grups (de grups de permutacions en aquella època) introduint-hi nous conceptes i resultats, motivats per la seva investigació sobre la resolució per radicals de les equacions de grau ≥ 5, reprenent així l'herència de Lagrange. Galois introdueix les classes mòdul un subgrup, el producte de grups, el concepte d'isomorfisme, etc. L'obra de Galois, que comença quan amb 16 anys llegeix els treballs de Lagrange i acaba 4 anys més tard després d'una joventut turbulenta i d'una mort tràgica, és un dels capítols més apassionants de les matemàtiques. (Llegiu *Obra d'Évariste Galois*, Institut d'Estudis Catalans, Monografies de la Secció de Ciències n.1 1984).

A partir de la segona meitat del segle 19 la teoria de grups es va configurant de mica en mica, primer de la mà d'Arthur Cayley (1821–1895), que va donar la definició abstracta de grup finit, i després amb Camille Jordan (1838–1922), amb el seu llibre *Mémoire sur les groupes de mouvements*, i amb Walter von Dyck (1856–1934), que considera ja els grups abstractes a finals del segle passat i introdueix, entre altres, el concepte de sistema de generadors.

III.10 Exercicis

1. Demostreu que un cicle d'ordre n no es pot posar mai com a producte de menys de $n-1$ transposicions.

2. Demostreu que $\mathcal{S}_n$ admet els següents sistemes de generadors:

   a) $(1,2),(1,3),\ldots,(1,n)$.

   b) $(1,2),(2,3),\ldots,(n-1,n)$.

   c) $(1,2,\ldots,n),(1,2)$.

3. Donada la permutació

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 7 & 8 & 9 & 4 & 5 & 2 & 1 & 6 \end{pmatrix},$$

   calculeu σ^{100}.

4. Sigui G un subgrup de $\mathcal{S}_n$ no contingut a $\mathcal{A}_n$. Demostreu que exactament la meitat de les permutacions de G són parelles.

5. Una permutació es diu *regular* si un cop descomposta en producte de cicles disjunts tots els cicles tenen el mateix ordre. Demostreu que una permutació és regular si i només si és una potència d'un cicle d'ordre màxim.

6. Demostreu que les arrels de l'equació $z^6-1=0$ amb el producte de $\mathbf{C}$ formen un grup cíclic. Trobeu els generadors i els subgrups.

7. Demostreu que l'aplicació

$$\begin{array}{ccc} \mathbf{R} & \longrightarrow & \mathbf{C} \\ t & \longmapsto & e^{2\pi i t} = \cos 2\pi t + i \sin 2\pi t \end{array}$$

   és un morfisme de grups i expliciteu el teorema d'isomorfisme.

8. Demostreu que, per a un grup G, les següents afirmacions són equivalents:

   a) G es abelià.

   b) $x \mapsto x^{-1}$ és un morfisme.

   c) $x \mapsto x^2$ és un morfisme.

9. Si G és un grup cíclic d'ordre n i r és un divisor de n, demostreu que G té com a màxim $r-1$ elements d'ordre r. Si r és primer, aleshores hi ha exactament $r-1$ elements d'ordre r.

10. Demostreu que en tot grup G d'ordre p^r, p primer, existeix almenys un subgrup d'ordre p.

11. Si G és un grup d'ordre n tal que tot element (diferent del neutre) té ordre 2, demostreu que n és una potència de 2.

12. Sigui G el grup generat per les matrius

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{i} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

amb el producte de matrius. Demostreu que G és un grup no abelià de 8 elements.

13. Sigui G un grup i $G' = \{x^2 \mid x \in G\}$. Demostreu:

a) Si G és abelià, G' és un subgrup normal de G. (Doneu un contraexemple en el cas G no abelià).

b) Si G és abelià, G/G' no té quadrats llevat del 0.

c) Si H és un subgrup de G que conté G', H és normal i G/H és un grup abelià.

d) Calculeu G' i G/G' en els casos següents: $\mathbf{Z}$, $\mathbf{Z}/(n)$, $\mathbf{Q}$, $\mathcal{S}_3$.

14. Es defineix el *centre* $\mathrm{Z}\,G$ d'un grup G com el conjunt dels elements que commuten amb tots els elements de G. Demostreu:

a) $\mathrm{Z}\,G$ és un subgrup normal de G.

b) Si G conté un únic element d'ordre 2, aquest pertany a $\mathrm{Z}\,G$.

c) Si H és un subgrup normal de G contingut a $\mathrm{Z}\,G$ i G/H és cíclic, aleshores G és abelià.

d) $\mathrm{Z}(\mathcal{S}_n) = \{e\}$.

15. Es defineix el *commutador* G' d'un grup G com el subgrup generat pels elements de la forma $xyx^{-1}y^{-1}$ amb $x, y \in G$:

$$G' = \langle xyx^{-1}y^{-1} \mid x, y \in G \rangle.$$

Demostreu:

a) G' és un subgrup normal de G.

b) $G_{\mathrm{ab}} = G/G'$ és un grup abelià (es diu l'*abelianitzat* de G).

c) Per a tot subgrup normal H de G que contingui G', G/H és abelià.

d) Si $f : G \longrightarrow A$ és un morfisme de G en un grup abelià A, existeix un únic morfisme $f' : G_{\mathrm{ab}} \longrightarrow A$ tal que el diagrama següent és commutatiu:

$$\begin{array}{ccc} & & f \\ G & \longrightarrow & A \\ \pi \downarrow & \nearrow_{f'} & \\ G_{\mathrm{ab}} & & \end{array}$$

on π és l'epimorfisme canònic que aplica cada element de G en la seva classe mòdul G'.

e) La propietat expressada a d) caracteritza el grup G_{ab}.

f) Determineu G' i G_{ab} per als grups que coneixeu.

III.11 Exercicis de programar

16. Producte de dues permutacions. (Indicació: per guardar una permutació $\sigma \in \mathcal{S}_n$ guardeu $\sigma(1), \ldots, \sigma(n)$).

17. Ordre d'una permutació. (Indicació: genereu $\sigma^2, \sigma^3, \ldots$ i combineu l'exercici anterior amb un petit dispositiu que reconegui quan una permutació és la identitat).

18. Descomposició d'una permutació en cicles disjunts. Signe. (Indicació: apliqueu el mètode de la proposició 2.2).

19. Subgrup generat per un conjunt de permutacions. (Indicació: prepareu primer una subrutina que donada una permutació us la compari amb totes les d'una llista $\tau_1, \ldots, \tau_k$ prèviament obtinguda).

Suggeriment: utilitzeu variables alfanumèriques per tal d'obtenir, a la llista $\tau_1, \ldots, \tau_k$, a més de les diferents permutacions, llurs expressions en funció dels generadors. Això permet estalviar operacions i conèixer l'expressió de qualsevol permutació del subgrup en funció dels generadors.

Capítol IV

Espais vectorials

IV.1 Definició i exemples

D'ara endavant, si no diem el contrari, K indicarà un cos commutatiu. Un *espai vectorial sobre K és un conjunt E no buit junt amb*

1. una operació $+$, que anomenarem *suma,* que compleix les següents propietats:
   - és associativa: $u+(v+w)=(u+v)+w \quad \forall u,v,w \in E$.
   - és commutativa: $u+v=v+u \quad \forall u,v \in E$.
   - hi ha un element $\vec{0}$ tal que $u+\vec{0}=u. \ \forall u \in E$
   - per a tot $u \in E$ hi ha un altre element, que es denota per $-u$, tal que $u+(-u)=\vec{0}$
2. una aplicació
$$\begin{array}{ccc} K \times E & \longrightarrow & E \\ (a,u) & \longmapsto & au \end{array}$$
que anomenarem *producte per elements de K*, que compleix
   - $a(u+v)=au+av \quad \forall a \in K, u,v \in E$
   - $(a+b)u=au+bu \quad \forall a,b \in K, u \in E$
   - $(ab)u=a(bu) \quad \forall a,b \in K, u \in E$
   - $1u=u \quad \forall u \in E$, on 1 és la unitat del cos K.

Observem que 1 diu que $(E,+)$ és un grup commutatiu (III.1). Els elements de E els anomenarem *vectors,* els de K *escalars.* Farem servir la notació $u-v$ per indicar $u+(-v)$.

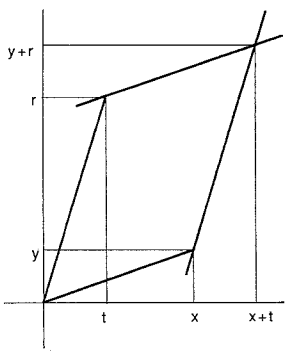
De la definició es dedueix fàcilment:

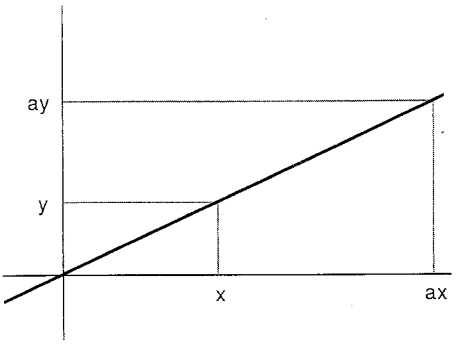
- $0v = \vec{0}$. En efecte, $0v = (0+0)v = 0v + 0v \Rightarrow 0v = \vec{0}$.
- $a\vec{0} = \vec{0}$. En efecte, $a\vec{0} = a(\vec{0}+\vec{0}) = a\vec{0} + a\vec{0} \Rightarrow a\vec{0} = \vec{0}$.
- $av = \vec{0} \Rightarrow a = 0$ o $v = \vec{0}$. En efecte, si $a \neq 0$, a té un invers a^{-1}. Llavors $v = 1v = (a^{-1}a)v = a^{-1}(av) = a^{-1}\vec{0} = \vec{0}$.
- $(-1)v = -v$. En efecte, $v + (-1)v = 1v + (-1)v = (1+(-1))v = 0v = \vec{0}$.

Exemples:

1. Sigui
$$\begin{cases} a_1^1 x^1 + \ldots + a_n^1 x^n = 0 \\ \ldots\ldots\ldots\ldots \\ a_1^m x^1 + \ldots + a_n^m x^n = 0 \end{cases}$$
un sistema homogeni de m equacions lineals amb n incògnites, i amb coeficients en un cos K. Una solució és una n-pla $(s^1, \ldots, s^n) \in K^n$ tal que en substituir $x^1, \ldots, x^n$ per aquests elements en el sistema totes les igualtats són certes. Si $(r^1, \ldots, r^n)$ és una altra solució, $(s^1+r^1, \ldots, s^n+r^n)$ tambe ho és; i si $a \in K$, $(as^1, \ldots, as^n)$ és també una solució. El conjunt de solucions del sistema donat, amb les dues operacions suma i producte per elements de K que acabem de definir, és un espai vectorial sobre K.

2. Considerem a $\mathbf{R}^2$ les dues operacions
$$(x,y) + (t,r) = (x+t, y+r)$$
$$a(x,y) = (ax, ay).$$
Amb aquestes operacions $\mathbf{R}^2$ és un espai vectorial sobre $\mathbf{R}$. Sovint els parells (x,y) de $\mathbf{R}^2$ es representen com a punts d'un pla. La suma que





hem definit coincideix amb la coneguda "llei del paral·lelogram", segons la qual es sumen les forces a Física. El producte per un $a \in \mathbf{R}$ coincideix, igualment, amb el producte d'una força per un $a \in \mathbf{R}$.

3. K^n amb les operacions

$$(x^1,\dots,x^n)+(y^1,\dots,y^n)=(x^1+y^1,\dots,x^n+y^n)$$
$$a(x^1,\dots,x^n)=(ax^1,\dots,ax^n)$$

és un espai vectorial sobre K. L'exemple 2 és un cas particular d'aquest.

4. K és un espai vectorial sobre ell mateix. El producte és el producte ordinari de K. $\mathbf{C}$ és un espai vectorial sobre $\mathbf{C}$. $\mathbf{C}$ és també un espai vectorial sobre $\mathbf{R}$, ja que existeix un producte d'elements de $\mathbf{R}$ per elements de $\mathbf{C}$ amb les propietats necessàries. I també és un espai vectorial sobre els racionals $\mathbf{Q}$, pel mateix motiu.

5. El conjunt de polinomis $K[x]$ és un espai vectorial sobre K amb les operacions usuals.

6. Anomenarem *matriu* $m \times n$ un quadre d'elements de K:

$$\begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix} \quad a_i^j \in K.$$

Designarem per $M_{m\times n}(K)$ el conjunt de les matrius $m \times n$ sobre K. En aquest conjunt definim una suma i un producte per elements de K de la següent manera:

$$\begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix} + \begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & & \vdots \\ b_1^m & \dots & b_n^m \end{pmatrix} = \begin{pmatrix} a_1^1+b_1^1 & \dots & a_n^1+b_n^1 \\ \vdots & & \vdots \\ a_1^m+b_1^m & \dots & a_n^m+b_n^m \end{pmatrix}$$

$$k\begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix} = \begin{pmatrix} ka_1^1 & \dots & ka_n^1 \\ \vdots & & \vdots \\ ka_1^m & \dots & ka_n^m \end{pmatrix}.$$

Amb aquestes operacions $M_{m\times n}(K)$ és un espai vectorial sobre K.

IV.2 Subespais vectorials

Sigui E un espai vectorial sobre K. Un subconjunt no buit $F \subset E$ es diu un *subespai vectorial de* E si

1. $u, v \in F \Rightarrow u+v \in F$
2. $u \in F,\ k \in K \Rightarrow ku \in F.$

Aquestes dues condicions ens diuen que les operacions de E permeten definir unes operacions a F. Observem que amb aquestes operacions F és automàticament un espai vectorial sobre K.

Exemple:

El conjunt de solucions d'un sistema homogeni amb n incògnites és un subespai vectorial de K^n (exemples 1 i 3 del §1).

Un vector u és *combinació lineal* dels vectors $v_1, \ldots, v_n$ si existeixen $a^1, \ldots, a^n \in K$ tals que

$$u = a^1 v_1 + \ldots + a^n v_n.$$

De les condicions 1 i 2 de la definició de subespai vectorial resulta que tota combinació lineal de vectors $v_1, \ldots, v_n$ de F és un vector de F. Suposem que S és un subconjunt qualsevol de E. Designem per $\langle S \rangle$ el conjunt de les combinacions lineals d'elements de S. Tot subespai vectorial F que contingui S haurà de contenir $\langle S \rangle$. Per altra banda $\langle S \rangle$ és, ell mateix, un subespai vectorial. Tenim doncs la següent

Proposició 2.1 *Si S és un subconjunt d'un espai vectorial E, el conjunt $\langle S \rangle$ és el més petit dels subespais vectorials de E que contenen S.*

Si $\langle S \rangle = F$ es diu que S *genera* F, que F està *generat* per S o que S és un *sistema de generadors* de F.□

Exemples:

1. $\mathbf{R}^2 = \langle (1,0), (0,1) \rangle$, ja que tot parell $(x,y) \in \mathbf{R}^2$ és de la forma

$$(x,y) = x(1,0) + y(0,1).$$

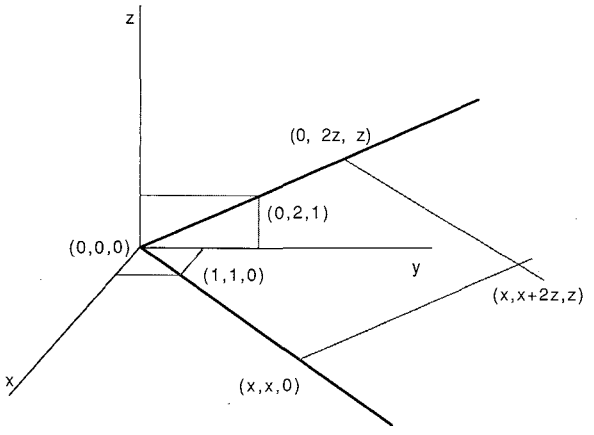
2. $K^n = \langle (1,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,\ldots,0,1) \rangle$.
3. $K[x] = \langle 1, x, x^2, \ldots, x^n, \ldots \rangle$.
4. $F = \{(x, 2z+x, z) \mid x, z \in \mathbf{R}\}$ és un subespai vectorial de $\mathbf{R}^3$. Els seus elements són de la forma

$$(x, 2z+x, z) = x(1,1,0) + z(0,2,1).$$

Resulta doncs que $F = \langle (1,1,0), (0,2,1) \rangle$. Representem els elements de R^3 com a punts de l'espai en la forma usual. Els elements de F són, en aquesta representació, els punts del pla que passa per $(0,0,0)$, $(1,1,0)$ i $(0,2,1)$.

IV.3 Bases d'un espai vectorial

Un conjunt de vectors S es diu *linealment independent* si tota combinació lineal de vectors de S nul·la té tots els coeficients nuls: $a^1 v_1 + \ldots + a^m v_m = \vec{0}, v_i \in S, i = 1, \ldots, m \Rightarrow a^1 = \ldots = a^m = 0$. En cas contrari direm que S és *linealment dependent* .



Proposició 3.1 *$v_1, \ldots, v_m$ són linealment dependents si i només si un d'ells és combinació lineal dels altres.*

DEMOSTRACIÓ: Si $v_1, \ldots, v_m$ són linealment dependents hi ha una combinació lineal

$$a^1 v_1 + \ldots + a^m v_m = \vec{0}$$

amb algun coeficient no nul. Podem suposar que $a^1 \neq 0$ (reordenant $v_1, \ldots, v_m$ si convé). Aleshores existeix $(a^1)^{-1}$ i

$$v_1 = -(a^1)^{-1} a^2 v_2 - (a^1)^{-1} a^3 v_3 - \ldots - (a^1)^{-1} a^m v_m.$$

Recíprocament,

$$v_1 = a^2 v_2 + \ldots + a^m v_m \Rightarrow 1 v_1 - a^2 v_2 - \ldots - a^m v_m = \vec{0}.$$

El primer coeficient no és nul i, per tant, $v_1, \ldots, v_m$ són linealment dependents. □

Exemples:

1. Un únic vector v és linealment independent si i només si $v \neq \vec{0}$.
2. A $\mathbf{R}^2$, $(1,0)$ i $(0,1)$ són linealment independents. En general, a K^n

   $$(1,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,\ldots,0,1)$$

   són linealment independents.
3. A $\mathbf{R}^3$, $(-1,1,0), (5,2,3), (0,7,3)$ són linealment dependents, ja que

   $$5(-1,1,0) + (5,2,3) - (0,7,3) = (0,0,0).$$

4. A $K[x]$, $S = \{1, x, x^2, \ldots, x^n, \ldots\}$ és linealment independent.

Una *base* d'un espai vectorial E és un sistema de generadors linealment independents.

Proposició 3.2 *$B \subset E$ és una base de E si i només si tot $u \in E$ s'expressa de manera única com una combinació lineal d'elements de B.*

DEMOSTRACIÓ: Demostrarem successivament els dos sentits de la implicació.

$\Rightarrow$) Donat $u \in E$, u és combinació d'elements de B ja que B genera E. Siguin $u = a^1v_1 + \ldots a^nv_n$ i $u = b^1u_1 + \ldots + b^mu_m$ dues expressions de u com a combinació lineal de vectors de B. Si un vector v_i no apareix en la segona expressió, afegim a aquesta el sumand $0v_i$; anàlogament, si un u_i no apareix en la primera expressió, afegim a aquesta el sumand $0u_i$. Obtenim així dues combinacions lineals dels mateixos vectors. Siguin

$$u = a^1v_1 + \ldots + a^rv_r = b^1v_1 + \ldots + b^rv_r.$$

Restant, obtenim

$$(a^1 - b^1)v_1 + \ldots + (a^r - b^r)v_r = \vec{0}.$$

Els vectors $v_1, \ldots, v_r$ són de B i, per tant, linealment independents. Els coeficients d'aquesta combinació lineal han d'ésser, doncs, nuls. Per tant,

$$a^1 = b^1, \ldots, a^r = b^r.$$

$\Leftarrow$) Tot vector de E s'expressa com a combinació lineal de vectors de B. És a dir, $\langle B \rangle = E$. Si $a^1v_1 + \ldots + a^mv_m = \vec{0}$ amb $v_i \in B$, $i = 1, \ldots, n$, com que també $0v_1 + \ldots + 0v_m = \vec{0}$ i l'expressió ha d'ésser única,

$$a^1 = 0, \ldots, a^n = 0.$$

B és, doncs, linealment independent. □

Observació:

Si S és linealment independent, S és base de $\langle S \rangle$.

Proposició 3.3 *Si S és linealment independent i $u \notin \langle S \rangle$, llavors $S \cup \{u\}$ és linealment independent.*

DEMOSTRACIÓ: Considerem

$$au + a^1v_1 + \ldots + a^mv_m = \vec{0}$$

amb $v_i \in S$, $i = 1, \ldots, m$. Si $a \neq 0$, existeix a^{-1} i

$$u = -a^{-1}a^1v_1 - \ldots - a^{-1}a^mv_m \in \langle S \rangle$$

en contra de la hipòtesi. Per tant, $a = 0$. Però llavors $a^1v_1 + \ldots + a^mv_m = \vec{0}$ i tots els vectors d'aquesta expressió són de S. Com que S és linealment independent, $a^1 = \ldots = a^m = 0$. □

Teorema 3.4 *Tot espai vectorial $E \neq \{\vec{0}\}$ generat per un nombre finit de vectors té una base finita.*

DEMOSTRACIÓ: Sigui $E = \langle v_1, \ldots, v_m \rangle$. Si els generadors són linealment independents, formen base. En cas contrari, n'hi ha un, sigui el v_i, que és combinació lineal dels altres. Aleshores

$$E = \langle v_1, \ldots, v_m \rangle = \langle v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m \rangle.$$

Si aquest nou conjunt de generadors és linealment independent, formen base. En cas contrari podem suprimir-ne un, obtenint un nou conjunt de generadors. Repetim el procés tantes vegades com sigui necessari, eliminant sempre aquells generadors que siguin combinació lineal dels altres. Arribarem d'aquesta manera a un conjunt linealment independent (és a dir, a una base), o a eliminar-los tots. En aquest segon cas seria $E = \{\vec{0}\}$. □

De fet aquesta demostració prova més del que diu l'enunciat. Prova que tot sistema de generadors conté una base.

Exemples:

1. Els vectors $(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)$ formen una base de K^n.
2. El conjunt $\{1, x, x^2, \ldots, x^n, \ldots\}$ és una base de $K[x]$.
3. Sigui E_i^j la matriu de $M_{m\times n}(K)$ formada per 0 a tot arreu, menys a la columna i , fila j en què apareix 1. El conjunt E_i^j ($i = 1,\ldots,n$ i $j = 1,\ldots,m$) és una base de $M_{m\times n}(K)$.
4. Els vectors $(1,1,0),(0,2,1)$ formen una base del subespai vectorial $F = \{(x, 2z + x, z) \mid x, z \in \mathbf{R}\}$. (Veure exemple 4 de l'apartat 2).

Exercici:

Proveu la veracitat de tots els exemples que acabem de donar.

Per les seves conseqüències, el següent teorema té una especial importància.

Teorema 3.5 (de Steinitz.) *Sigui $u_1, \ldots, u_n$ una base de E i siguin $v_1, \ldots, v_m$ vectors linealment independents. Aleshores es poden substituir m vectors de la base $u_1, \ldots, u_n$ per $v_1, \ldots, v_m$ obtenint una nova base. En particular, $m \leq n$.*

DEMOSTRACIÓ: Es tracta d'introduir un per un els $v_1, \ldots, v_m$ en substitució de vectors de la base donada.

1. Introducció de v_1. Per ser $u_1, \ldots, u_n$ una base tindrem

$$v_1 = \sum_{i=1}^{n} a^i u_i;$$

v_1 no és nul, i per tant un dels coeficients a^i és no nul. Podem suposar, reordenant si convé la base $u_1, \ldots, u_n$, que $a^1 \neq 0$. Llavors

$$u_1 = (a^1)^{-1}v_1 - \sum_{i=2}^{n}(a^1)^{-1}a^i u_i.$$

Aquesta expressió ens diu que

$$\langle u_1, u_2, \ldots, u_n\rangle = \langle v_1, u_2, \ldots, u_n\rangle.$$

A més a més, $v_1, u_2, \ldots, u_n$ són linealment independents. En efecte,

$$b^1v_1 + b^2u_2 + \ldots b^nu_n = \vec{0} \Rightarrow b^1(\sum_{i=1}^{n} a^iu_i) + b^2u_2 + \ldots + b^nu_n = \vec{0}$$

$$\Rightarrow \quad b^1a^1u_1 + \sum_{i=2}^{n}(b^1a^i + b^i)u_i = \vec{0} \Rightarrow b^1a^1 = 0, b^1a^i + b^i = 0, i \geq 2,$$

ja que $u_1, \ldots, u_n$ és una base. Però $a^1 \neq 0$. Per tant, $b^1 = 0$ i $b^i = 0$ $(i = 2, \ldots, n)$. Així, doncs, $v_1, u_2, \ldots, u_n$ formen una base de E.

2. Suposem que ja hem substituït h vectors de la base per $v_1, \ldots, v_h$. Reordenant si convé la base $u_1, \ldots, u_n$ podem suposar que hem substituït els h primers, i tenim que

$$v_1, \ldots, v_h, u_{h+1}, \ldots, u_n$$

és una base de E. Procedim igual que a 1 i posem v_{h+1} com combinació lineal dels vectors d'aquesta base:

$$v_{h+1} = \sum_{i=1}^{h} a^iv_i + \sum_{i=h+1}^{n} a^iu_i.$$

1 ens diu que podem substituir per v_{h+1} qualsevol vector que, en aquesta expressió, tingui coeficient no nul. Tot queda reduït, doncs, a veure que un dels coeficients del segon sumatori no és nul. Però, en efecte, si $a^{h+1} = \ldots = a^n = 0$, aleshores v_{h+1} seria combinació lineal de $v_1, \ldots, v_h$ i això no és cert, ja que tots els $v_1, \ldots, v_m$ són linealment independents. □

Corol·lari 3.6 *Si l'espai vectorial E té una base finita, totes les bases de E tenen el mateix nombre de vectors.*

DEMOSTRACIÓ: Siguin $u_1, \ldots, u_n$ i $\{v_j \mid j \in J\}$ dues bases de E. De (3.5) es dedueix que tota família finita $v_{j_1}, \ldots, v_{j_k}$ satisfà $k \leq n$. Per tant, J ha de ser finit. Llavors, si J té m elements, tenim $m \leq n$ i també $n \leq m$ (3.5). Per tant, $m = n$.□

La *dimensió* d'un espai vectorial E sobre un cos K és el nombre d'elements de les seves bases, si són finites. En cas contrari, direm que E és de dimensió infinita.

Corol·lari 3.7 *La dimensió d'un espai coincideix amb el nombre màxim d'elements linealment independents, i també amb el nombre mínim de generadors.*

DEMOSTRACIÓ: La primera afirmació és conseqüència immediata del teorema de Steinitz. La segona resulta de la demostració de (3.4). □

Corol·lari 3.8 *Tot conjunt de vectors linealment independents pot completar-se fins a obtenir una base.*□

Exemples:

1. K^n és de dimensió n sobre K.
2. $K[x]$ és de dimensió infinita sobre K.
3. $M_{m\times n}(K)$ és de dimensió nm sobre K.
4. Els complexos $\mathbf{C}$ són un espai vectorial sobre $\mathbf{C}$ de dimensió 1, i un espai vectorial sobre $\mathbf{R}$ de dimensió 2. En el segon cas $\{1, i\}$ és una base.

Proposició 3.9 *Sigui F un subespai de l'espai vectorial E. Si la dimensió de E és finita, la de F també i*

$$\dim F \leq \dim E.$$

A més a més,

$$\dim F = \dim E \Leftrightarrow F = E.$$

DEMOSTRACIÓ: Si $F = \{\vec{0}\}$ no hi ha res a dir. En cas contrari, sigui $\vec{0} \neq v_1 \in F$; si $F = \langle v_1 \rangle$, v_1 és base. En cas contrari, sigui $v_2 \in F$, $v_2 \notin \langle v_1 \rangle$; si $F = \langle v_1, v_2 \rangle$, $\{v_1, v_2\}$ formen base per (3.3). En cas contrari, sigui $v_3 \notin F$, $v_3 \notin \langle v_1, v_2 \rangle \ldots$ Per (3.7) aquest procés ha d'acabar. Haurem trobat, aleshores, una base de F que tindrà com a màxim n elements (on n és la dimensió de E). Si $\dim F = n$ i $v_1, \ldots, v_n$ és una base de F, per (3.5) també és una base de E. Llavors

$$F = \langle v_1, \ldots, v_n \rangle = E. \quad \square$$

IV.4 Fórmula de Grassmann. Suma directa de subespais.

Sigui E un espai vectorial i F, G dos subespais de E.

Proposició 4.1 *$F \cap G$ és un subespai vectorial de E.*

Exercici:

Demostreu (4.1)

En general $F \cup G$ no és un subespai vectorial de E. El motiu és que la suma d'un vector de F i un vector de G pot no ser ni a F ni a G.

Exemple:

Considerem els subespais $F = \{(x,0) \mid x \in \mathbf{R}\}$ i $G = \{(0,y) \mid y \in \mathbf{R}\}$ de $\mathbf{R}^2$. La suma $(1,0)+(0,1)$ no és ni a F ni a G.

Per evitar el treball amb conjunts que no són subespais vectorials, normalment considerem, en comptes de la unió $F \cup G$, el subespai vectorial generat per aquesta unió. Aquest subespai és precisament

$$\{u+v \mid u \in F, v \in G\}.$$

En efecte, és fàcil veure que aquest és el més petit dels subespais que contenen F i G. En direm *suma* de F i G i el designarem per $F+G$.

Teorema 4.2 (Fórmula de Grassmann) *Siguin F i G dos subespais vectorials de E i suposem que la dimensió de E és finita. Aleshores F, G, $F \cap G$ i $F+G$ són tots de dimensió finita i*

$$\dim F + \dim G = \dim(F+G) + \dim(F \cap G).$$

DEMOSTRACIÓ: Per (3.9) tots ells són de dimensió finita. Sigui $u_1, \ldots, u_m$ una base de $F \cap G$. Podem completar aquesta base fins a obtenir una base de F i una base de G (per (3.8)): $u_1, \ldots, u_m, u_{m+1}, \ldots, u_r$ base de F, $u_1, \ldots, u_m, v_{m+1}, \ldots, v_s$ base de G. Tot vector de la forma $u+v$ amb $u \in F$ i $v \in G$ serà, doncs, combinació lineal de $u_1, \ldots, u_m, u_{m+1}, \ldots, u_r, v_{m+1}, \ldots, v_s$. Si demostrem que tots aquests vectors són linealment independents haurem obtingut una base de $F+G$ amb un nombre de vectors que demostra la igualtat de l'enunciat. Sigui doncs

$$\sum_{i=1}^{r} a^i u_i + \sum_{i=m+1}^{s} b^i v_i = \vec{0}.$$

Aleshores,

$$\sum_{i=1}^{r} a^i u_i = -\sum_{i=m+1}^{s} b^i v_i \in F \cap G,$$

d'on $\sum_{i=m+1}^{s} b^i v_i = \sum_{j=1}^{m} c^j u_j$, és a dir, $\sum_{j=1}^{m} c^j u_j + \sum_{i=m+1}^{s} b^i v_i = \vec{0}$ i, com que $u_1, \ldots, u_m, v_{m+1}, \ldots, v_s$ són una base de G, $c^j = 0$ $(j = 1, \ldots, m)$ i $b^i = 0$, $(i = m+1, \ldots, s)$. Per tant, $\sum_{i=1}^{r} a^i u_i = \vec{0}$ i, com que $u_1, \ldots, u_r$ són una base de F,

$$a^i = 0, i = 1, \ldots, r.$$

És a dir, en la combinació lineal inicial tots els coeficients són 0. □

Si $F \cap G = \{\vec{0}\}$ direm que la suma $F + G$ és una *suma directa* i escriurem

$$F \oplus G.$$

El teorema 4.2 ens diu que la dimensió de $F \oplus G$ és la suma de les dimensions dels dos subespais F i G. La proposició que segueix dóna una caracterització de les sumes directes.

Proposició 4.3 *La suma $F + G$ és directa si i només si l'expressió d'un vector de $F + G$ com a suma d'un vector de F i un vector de G és única.*

DEMOSTRACIÓ: Provem la implicació $\Rightarrow$: si tenim dues expressions $u+v = u_1+v_1$ amb $u, u_1 \in F$ i $v, v_1 \in G$, aleshores $u - u_1 = v_1 - v \in F \cap G = \vec{0}$, d'on $u - u_1 = v_1 - v = \vec{0}$ i, per tant, $u = u_1$, $v = v_1$.

Provem la implicació $\Leftarrow$: si $w \in F \cap G$, resulta que $w + \vec{0} = \vec{0} + w$ són dues expressions del mateix vector de $F + G$. Les dues expressions han de coincidir. Per tant, $w = \vec{0}$. □

Proposició 4.4 *Si la dimensió de E és finita, per a tot subespai F hi ha un altre subespai G tal que $E = F \oplus G$.*

DEMOSTRACIÓ: Sigui $u_1, \ldots, u_m$ una base de F. Completem-la a una base de E, $u_1, \ldots, u_m, u_{m+1}, \ldots, u_n$ (3.8). El subespai $G = \langle u_{m+1}, \ldots, u_n \rangle$ compleix l'enunciat de la proposició. □

El subespai esmentat a (4.4) es diu un *complementari de F*. F té, en general, molts complementaris.

Tot el que hem fet en aquest apartat pot ésser generalitzat a un nombre finit de subespais $F_1, \ldots, F_k$. Per a la intersecció, la generalització de (4.1) és clara: $F_1 \cap \ldots \cap F_k$ és sempre un subespai vectorial. La unió $F_1 \cup \ldots \cup F_k$ no és sempre un subespai vectorial; definim la *suma*

$$F_1 + \cdots + F_k$$

com el subespai generat per $F_1 \cup \ldots \cup F_k$. Resulta que

$$F_1 + \cdots + F_k = \{v_1 + \cdots + v_k \mid v_i \in F_i, i = 1, \ldots, k\}.$$

La generalització de suma directa presenta més dificultats. La forma correcta de fer-ho es pel camí de (4.3). Així, direm que la suma $F_1 + \cdots + F_k$ és *suma directa* i escriurem

$$F_1 \oplus \cdots \oplus F_k$$

si l'expressió de tot vector de $F_1 + \cdots + F_k$ com a suma de vectors de $F_1, \ldots, F_k$ és única.

Exercici:

Demostreu que $F_1 + \cdots + F_k$ és una suma directa si i només si, per a tot $i = 1, \ldots, k$, $F_i \cap (F_1 + \cdots + F_{i-1} + F_{i+1} + \cdots + F_k) = \{\vec{0}\}$.

IV.5 Suma directa d'espais vectorials

Siguin E i F dos espais vectorials sobre el mateix cos K. Anomenarem *suma directa* de E i F el conjunt $E \times F$ amb les operacions

$$(u, v) + (u_1, v_1) = (u + u_1, v + v_1)$$

$$k(u, v) = (ku, kv),$$

on $u, u_1 \in E$, $v, v_1 \in F$ i $k \in K$. Amb aquestes operacions $E \times F$ és un espai vectorial que designarem per

$$E \oplus F.$$

Exemple:

$$K^n = \overbrace{K \oplus \cdots \oplus K}^{n}.$$

Proposició 5.1 *Si E i F són de dimensió finita, $E \oplus F$ també i* $\dim E \oplus F = \dim E + \dim F$.

DEMOSTRACIÓ: Sigui $u_1, \ldots, u_n$ una base de E i $v_1, \ldots, v_m$ una base de F. Llavors $(u_1, \vec{0}), \ldots, (u_n, \vec{0}), (\vec{0}, v_1), \ldots, (\vec{0}, v_m)$ és una base de $E \oplus F$. En efecte: aquests vectors generen $E \oplus F$, ja que tot $(u, v) \in E$ és

$$(u, v) = (u, \vec{0}) + (\vec{0}, v) = (\sum_{i=1}^{n} a^i u_i, \vec{0}) + (\vec{0}, \sum_{j=1}^{m} b^j v_j) = \sum_{i=1}^{n} a^i (u_i, \vec{0}) + \sum_{j=1}^{m} b^j (\vec{0}, v_j),$$

i són linealment independents, ja que si

$$\sum_{i=1}^{n} a^i (u_i, \vec{0}) + \sum_{j=1}^{m} b^j (\vec{0}, v_j) = (\vec{0}, \vec{0})$$

aleshores

$$(\sum_{i=1}^{n} a^i u_i, \sum_{j=1}^{m} b^j v_j) = (\vec{0}, \vec{0}),$$

que implica $\sum_{i=1}^{n} a^i u_i = \vec{0}$, d'on $a^i = 0 \; \forall i$, i $\sum_{j=1}^{m} b^j v_j = \vec{0}$, d'on $b^j = 0 \; \forall j$. □

Tant el nom com la proposició 5.1 suggereixen una relació entre la suma directa d'espais i la suma directa de subespais vectorials d'un espai E. D'aquesta relació en parlarem a V.1. Exemple 6.

IV.6 Espai vectorial quocient.

Sigui E un espai vectorial i F un subespai vectorial de E. Direm que $u, v \in E$ estan *relacionats mòdul* F si $u - v \in F$. Aquesta relació és d'equivalència (I.4) i es pot formar el corresponent conjunt quocient que designarem per E/F.

La classe $[u]$ d'un vector $u \in E$ és el conjunt $\{u + v \mid v \in F\}$ i la denotarem també per $u + F$.

Si u i v són equivalents mòdul F a u_1 i v_1, respectivament, llavors $(u + v)$ i $(u_1 + v_1)$ son equivalents mòdul F. Podem, doncs, definir una operació suma

$$[u] + [v] = [u + v]$$

que no depèn dels representants. Anàlogament l'operació

$$k[u] = [ku] \qquad k \in K$$

no depèn dels representants. Aquestes dues operacions fan de E/F un espai vectorial sobre K.

Proposició 6.1 *Si E és de dimensió finita, E/F també ho és i*

$$\dim E/F = \dim E - \dim F.$$

DEMOSTRACIÓ: Completem una base de F , $u_1, \ldots, u_m$, fins a obtenir una base de E : $u_1, \ldots, u_m, u_{m+1}, \ldots, u_n$. Per a $i = 1, \ldots, m$, $[u_i] = [\vec{0}]$; per a $i > m$, les classes

$$[u_{m+1}], \ldots, [u_n]$$

constitueixen una base de E/F; en efecte, tota classe $[u]$ d'un vector $u = \sum_{i=1}^{n} a^i u_i$ es pot escriure com $[u] = \sum_{i=m+1}^{n} a^i [u_i]$ i, per tant, aquestes classes generen E/F. Per veure que són linealment independents considerem

$$\sum_{i=m+1}^{n} a^i [u_i] = [\vec{0}].$$

Aleshores $\left[\sum_{i=m+1}^{n} a^i u_i\right] = [\vec{0}]$, d'on $\sum_{i=m+1}^{n} a^i u_i \in F$. Aquest vector es pot expressar, doncs, en la base de F: $\sum_{i=m+1}^{n} a^i u_i = \sum_{j=1}^{m} b^j u_j$, és a dir,

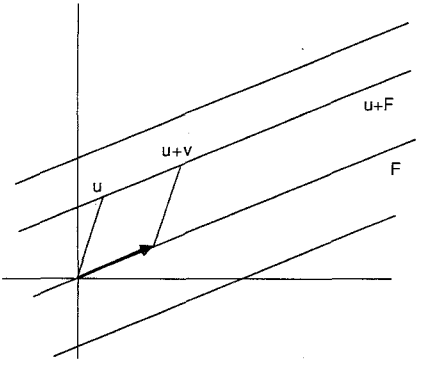
$$\sum_{j=1}^{m} b^j u_j - \sum_{i=m+1}^{n} a^i u_i = \vec{0}.$$

Els vectors que apareixen aquí formen una base de E i, per tant, els coeficients s'han d'anul·lar. En particular, $a^i = 0$ per a $i = m+1, \ldots, n$.

Hem obtingut una base de E/F, demostrant, així, la igualtat de l'enunciat. □

Exemples:

1. Considerem un subespai $F = \langle v \rangle$ de $\mathbf{R}^2$. Si representem els elements de $\mathbf{R}^2$ com a punts del pla, els vectors de F corresponen als punts d'una recta que passa per l'origen. Cada una de les classes $u + F$ correspon, llavors, a punts d'una recta paral·lela a F que passa per u. El conjunt $\mathbf{R}^2/F$ és, en aquest cas, el conjunt de rectes paral·leles a F. Anàlogament si $F = \langle v \rangle \subset \mathbf{R}^3$, $\mathbf{R}^3/F$ és el conjunt de rectes de $\mathbf{R}^3$ paral·leles a una recta F que passa per l'origen. Si $F = \langle u, v \rangle \subset \mathbf{R}^3$ és de dimensió 2, F correspon a un pla que passa per l'origen i $\mathbf{R}^3/F$ és el conjunt de plans paral·lels a F.



2. Designem per $\mathbf{R}_n[x]$ els polinomis de grau $\leq n$ de $\mathbf{R}[x]$. Dos polinomis $p(x) = p_0 + p_1 x + \cdots + p_r x^r$ i $q(x) = q_0 + q_1 x + \cdots + q_s x^s$ determinen la mateixa classe mòdul $\mathbf{R}_n[x]$ si i només si $p_i = q_i$ per a $i > n$.

IV.7 Coordenades

Fixada una base $u_1, \ldots, u_n$ en un espai vectorial E, l'expressió de cada vector v de E

$$v = a^1 u_1 + \cdots + a^n u_n$$

és única (3.2). Direm llavors que $(a^1, \ldots, a^n)$ són les *coordenades de v en la base* $u_1, \ldots, u_n$. Naturalment, les coordenades de v en una altra base $e_1, \ldots, e_n$ seran unes altres:

$$v = b^1 e_1 + \cdots + b^n e_n.$$

Quina relació hi ha entre $(a^1, \ldots, a^n)$ i $(b^1, \ldots, b^n)$? Suposem que

$$e_i = \sum_{j=1}^{n} p_i^j u_j, \quad i = 1, \ldots, n.$$

Tenim

$$v = \sum_{i=1}^{n} b^i e_i = \sum_{i=1}^{n} b^i (\sum_{j=1}^{n} p_i^j u_j) = \sum_{j=1}^{n} (\sum_{i=1}^{n} b^i p_i^j) u_j,$$

d'on

$$a^j = \sum_{i=1}^{n} b^i p_i^j, \quad j = 1, \ldots, n.$$

Podem escriure aquestes igualtats d'una manera molt més còmoda fent servir el producte de matrius. Concretament: si $A = (a_i^j)$ és una matriu de n columnes i m files i $B = (b_i^j)$ és una matriu de r columnes i n files, el *producte de A per B* és una matriu $C = (c_i^j)$

$$C = AB$$

de r columnes i m files, en la qual

$$c_i^j = \sum_{k=1}^{n} a_k^j b_i^k.$$

Considerem en el nostre cas les matrius

$$P = \begin{pmatrix} p_1^1 & \ldots & p_n^1 \\ \vdots & & \vdots \\ p_1^n & \ldots & p_n^n \end{pmatrix} \quad A = \begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix} \quad B = \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix}.$$

Les igualtats que relacionen les coordenades de V en una i altra base es poden resumir així:

$$A = PB.$$

La matriu P es diu la *matriu del canvi de base.*

Efectuem ara un nou canvi de base. Sigui $v_1, \ldots, v_n$ una nova base:

$$v_j = \sum_{i=1}^{n} q_j^i e_i \qquad j = 1, \ldots, n.$$

$Q = (q_j^i)$ és la matriu del nou canvi de base. Quina relació hi ha entre les matrius P i Q d'aquests dos canvis de base successius i la matriu del canvi directe de la base $u_1, \ldots, u_n$ a la base $v_1, \ldots, v_n$? Només cal trobar les coordenades dels vectors v_j en la base $u_1, \ldots, u_n$:

$$v_j = \sum_{i=1}^{n} q_j^i e_i = \sum_{i=1}^{n} q_j^i \left(\sum_{k=1}^{n} p_i^k u_k \right) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} p_i^k q_j^i \right) u_k.$$

Aquests coeficients són, precisament, els de la matriu PQ; aquesta és, doncs, la matriu del canvi directe.

Què passa si, en aquests dos canvis successius, la tercera base torna a ser la primera $v_1 = u_1, \ldots, v_n = u_n$? El canvi directe té, evidentment, la matriu

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Aquesta matriu la designarem per I i en direm *matriu identitat*, ja que compleix

$$MI = IM = M$$

per a tota matriu $n \times n$ M. En aquest cas resulta doncs que

$$PQ = I.$$

Anàlogament, si fem primer el canvi de $e_1, \ldots, e_n$ a $u_1, \ldots, u_n$ de matriu Q , i després el canvi de $u_1, \ldots, u_n$ a $e_1, \ldots, e_n$ de matriu P, obtindrem

$$QP = I.$$

Dues matrius, el producte de les quals, en qualsevol ordre, es sempre I es diuen *inverses* una de l'altra. P i Q són inverses una de l'altra. Escriurem

$$P = Q^{-1} \quad \text{i} \quad Q = P^{-1}.$$

Hem demostrat, doncs, la

Proposició 7.1 *Les matrius dels canvis de base són sempre invertibles.* □

IV.8 Nota històrica

Les primeres idees sobre coordenades són de René Descartes (1596–1650), que introdueix la notació $x^1, x^2, x^3, \ldots$ i la regla dels signes per a polinomis en el *Discours de la méthode*, "... pour bien conduire la raison et chercher la vérité dans les sciences". La consideració de conceptes en dimensió n, així com la suma de vectors (en identificar $\mathbf{R}^2$ amb $\mathbf{C}$ en la demostració del teorema fonamental de l'àlgebra) són degudes a Carl Friedrich Gauss (1777–1855). Arthur Cayley (1821–1895) i Hermann G. Grassmann (1809–1877) (aquest darrer, un mestre d'escola sense formació científica) ja utilitzen els espais vectorials (que no serien definits axiomàticament fins el 1888 per Giuseppe Peano (1858–1932)). Grassmann introdueix els conceptes de subespai, generadors, dimensió, suma, intersecció, així com les fórmules per al canvi de coordenades. El concepte i el nom de matriu fou introduït per James Joseph Sylvester (1814–1897) l'any 1850 amb posterioritat a l'estudi dels determinants (veure Cap.VI). (Llegiu "Els orígens físics de l'anàlisi vectorial" dins *El desenvolupament de les matemàtiques al segle 19*, Institut d'Estudis Catalans, Arxius de la Secció de Ciències, LXXV, 1984).

IV.9 Exercicis

1. Demostreu que el conjunt E de les aplicacions $f : \mathbf{R} \longrightarrow \mathbf{R}$ es pot dotar de manera natural d'una estructura d'espai vectorial real.

a) Quines de les següents famílies d'elements de E són linealment independents?

(a) $\sin x$, $\cos x$, 1.
(b) e^x, e^{x+2}.
(c) $2, x+2, x^2$.
(d) 0, 1, $x+1$.

b) Expresseu (si és possible) els següents elements de E com a combinació lineal de les famílies corresponents:

(a) $\sin x$; 1, x ,x^2, ...
(b) x^2+x-1; $1, x-1, (x-1)^2$.
(c) 1; $x+1, x^2-1, x^3+1$.
(d) 0; $(x-1)^2, x, x^2+2, \sqrt{3}$.

c) Quins dels següents subconjunts de E són subespais vectorials?

$$E_1 = \{f \in E \mid f(-x) = f(x)\ \forall x \in R\}$$
$$E_2 = \{f \in E \mid f(-x) = -f(x)\ \forall x \in R\}$$
$$E_3 = \{f \in E \mid f \text{ és continua}\}$$
$$E_4 = \{f \in E \mid f(0) = f(1)\}$$
$$E_5 = \{f \in E \mid \ f \text{ dues vegades derivable i } f'' - f' + f = 0\}$$

2. Sigui $E = M_{n\times m}(R)$. Quins dels següents subconjunts de E són subespais vectorials?

$$E_1 = \{A \in E \mid a_1^1 = 0\}$$
$$E_2 = \{A \in E \mid a_1^1 + a_1^2 = 0\}$$
$$E_3 = \{A \in E \mid \textstyle\sum_{i=1}^{n} a_i^i = 0,\ n = m\}$$
$$E_4 = \{A \in E \mid a_j^i = a_i^j\ \forall i, j,\ n = m\}$$
$$E_5 = \{A \in E \mid a_i^j = a_r^s\ \forall i, j, r, s\}$$
$$E_6 = \{A \in E \mid \textstyle\sum_{j=1}^{m} a_i^j = 2\pi,\ i = 1, \ldots, n\}.$$

3. Sigui $p(x)$ un polinomi de $K[x]$ de grau n. Demostreu que $K[x]/(p(x))$ té una estructura natural d'espai vectorial per a la qual $[1], [x], \ldots, [x^{n-1}]$ forma una base.

4. Demostreu que el conjunt de les successions d'elements d'un cos K amb les operacions

$$(x_1, x_2, \ldots, x_n, \ldots) + (y_1, y_2, \ldots, y_n, \ldots) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n, \ldots)$$

$$k(x_1, x_2, \ldots, x_n, \ldots) = (kx_1, kx_2, \ldots, kx_n, \ldots)$$

és un espai vectorial sobre K. Considerem el conjunt S de les successions amb tots els elements 0 llevat d'un que valgui 1. És S linealment independent? És S una base?

5. Sigui F el subespai de les successions (a_n) tals que $a_k = a_{k-1} + 2a_{k-2}$. Trobeu les progressions geomètriques contingudes a F i comproveu que existeixen bases de F formades per progressions geomètriques. Expresseu qualsevol altra successió de F com a combinació lineal d'una d'aquestes bases i trobeu així el terme general de les successions de F.

6. Sigui $K_n[x]$ el conjunt dels polinomis de grau $\leq n$ junt amb el 0. Donat $0 \neq p(x) \in K_n[x]$, designem per $p'(x), p''(x), \ldots, p^{(n)}(x)$ les seves derivades. Demostreu que $p(x), p'(x), p''(x), \ldots, p^{(n)}(x)$ formen una base de $K_n[x]$. Escriviu $1, x, x^2, \ldots, x^n$ com combinaci" lineal d'aquesta base, en el cas $p(x) = 1 + x + x^2 + \cdots + x^n$.

7. Donada una matriu $A = (a_i^j) \in M_{n\times n}(K)$ anomenarem transposada de A a la matriu $A^t = (b_i^j)$ tal que $b_i^j = a_j^i$, per a tot i, j. Considerem els conjunts

$$S = \{A \in M_{n\times n}(K) \mid A = A^t\} \quad \text{(Matrius simètriques)}$$
$$H = \{A \in M_{n\times n}(K) \mid A = -A^t\} \quad \text{(Matrius hemisimètriques).}$$

Demostreu que S i H són subespais vectorials de $M_{n\times n}(K)$ i que, si a K $2 = 1 + 1 \neq 0$, llavors $M_{n\times n}(K) = S \otimes M$. Què passa si $K = \mathbf{Z}/(2)$?

8. Trobeu un sistema homogeni d'equacions lineals que tingui com a solucions els elements de $F = \langle (1,1,1,1), (2,1,0,3) \rangle$.

9. Trobeu una base de l'espai vectorial de les solucions del sistema de tres equacions: $x - y = 0$, $2x + y + z = 0$, $x + y - z = 0$. Suposeu que els coeficients són elements de: a) $\mathbf{R}$; b) $\mathbf{Z}/(2)$; c) $\mathbf{Z}/(5)$.

10. Considereu les inclusions $\mathbf{Q} \subset \mathbf{Q}(\sqrt[3]{5}) \subset \mathbf{R} \subset \mathbf{C}$, (veure l'exemple de II.7). Demostreu que cada un d'aquests cossos és un espai vectorial sobre l'anterior i determineu-ne les dimensions.

11. Siguin $M = \{\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \mid a, b, c \in \mathbf{R}\}$ i $M' = \{\begin{pmatrix} a-b & 2a \\ b & a+2b \end{pmatrix} \mid a, b \in \mathbf{R}\}$. Trobeu una base de $M_{2\times 2}(\mathbf{R})/M$ i de $M_{2\times 2}(\mathbf{R})/M'$.

12. Siguin F, G, H subespais de l'espai vectorial E. Proveu o doneu contraexemples de les afirmacions següents:

a) $F \cap (G + H) = (F \cap G) + (F \cap H)$.

b) $F + (G \cap H) = (F + G) \cap (F + H)$.

c) $\dim(F \cap (G + H)) = \dim(F \cap G) + \dim(F \cap H) + \dim(F \cap H \cap G)$.

13. Quines condicions han de complir a, b, c perquè els vectors de $\mathbf{R}^3$ (a, a^2, a^3), (b, b^2, b^3), (c, c^2, c^3) siguin linealment independents?

14. Considerem a $\mathbf{R}^4$ els subespais $F = \langle a, b, c\rangle$ i $G = \langle d, e\rangle$ on $a = (1, 2, 3, 4)$, $b = (2, 2, 2, 6)$, $c = (0, 2, 4, 4)$, $d = (1, 0, -1, 2)$ i $e = (2, 3, 0, 1)$.

a) Determineu les dimensions de $F, G, F \cap G$ i $F + G$ i doneu una base de cada subespai.

b) Doneu sengles bases de $\mathbf{R}^4/F$ i de $\mathbf{R}^4/G$.

IV.10 Exercicis de programar

15. Prepareu un programa per calcular, sobre $\mathbf{R}$,

a) la suma de dues matrius;

b) el producte d'una matriu per un escalar;

c) el producte de dues matrius.

16. Modifiqueu el programa anterior de manera que en lloc de treballar amb $K = \mathbf{R}$ ho faci amb

a) $K = \mathbf{C}$;

b) $K = \mathbf{Z}/(p)$.

Nota:

Convé preparar aquests programes en forma de subrutines que puguin ésser utilitzades quan convingui dins de programes més grans.

17. Descomposició *LU*

Sigui $A = (a_i^j) \in M_{n\times n}(\mathbf{R})$ una matriu donada. Prepareu un programa que descompongui $A = LU$, on $L = (l_i^j)$ és una matriu triangular inferior ($l_i^j = 0 \quad \forall i > j$) amb uns a la diagonal principal ($l_i^i = 1 \quad i = 1, \ldots, n$) i $U = (u_i^j)$ és una matriu triangular superior ($u_i^j = 0 \ \forall i < j$). (Indicació: efectuant el producte de les matrius L i U obtenim unes fórmules, que són les que s'han de programar). (Veure Cap VI, Exercicis de programació, per a una ampliació d'aquest exercici).

L'exercici següent utilitza l'exercici VII.12.

18. Si $e_1, \ldots, e_m$ és una família qualsevol de vectors de $\mathbf{R}^n$, prepareu un programa que permeti extreure'n una família linealment independent i completar-la a una base de $\mathbf{R}^n$ amb vectors de la base canònica. (Indicació: seguiu els passos de la demostració del teorema de Steinitz).

Capítol V

Aplicacions lineals

V.1 Definició i exemples

El capítol anterior ha estat dedicat a l'estudi d'unes estructures, els espais vectorials. En aquest estudiarem aplicacions entre aquestes estructures, les aplicacions lineals. És lògic pensar que les aplicacions interessants entre espais vectorials són aquelles que respecten l'estructura d'espai vectorial. Un espai vectorial és un conjunt amb dues operacions; una aplicació entre els conjunts que "conserva" les dues operacions és una "aplicació que respecta l'estructura vectorial".

Siguin E i F dos espais vectorials sobre el mateix cos K. Una aplicació

$$f : E \longrightarrow F$$

es diu una *aplicació lineal* si per a tot $u, v \in E$ i tot $a \in K$,

$$\begin{aligned} f(u+v) &= f(u)+f(v) \\ f(au) &= af(u). \end{aligned}$$

Si f és lineal, es compleix

- $f(au+bv) = af(u)+bf(v)$, i, en general,
- $f\left(\sum_{i=1}^{m} a^i u_i\right) = \sum_{i=1}^{m} a^i f(u_i)$.
- $f(\vec{0}) = \vec{0}$, ja que de $\vec{0} = 0v$ es dedueix $f(\vec{0}) = 0f(v) = \vec{0}$.
- $f(-v) = -f(v)$, ja que $f(v)+f(-v) = f(v+(-v)) = f(\vec{0}) = \vec{0}$.
- Si $f : E \longrightarrow F$ i $g : F \longrightarrow G$ són lineals, $g \circ f : E \longrightarrow G$ és lineal.

Exemples:

1. Sigui E un espai vectorial sobre K i $u_1, \ldots, u_n$ una base de E. L'aplicació

$$\begin{array}{ccc} E & \to & K^n \\ v & \mapsto & (a^1, \ldots, a^n) \end{array}$$

que fa correspondre a cada vector les seves coordenades en la base donada és lineal.

2. Sigui $E_1 \oplus E_2$ la suma directa dels espais vectorials E_1 i E_2. Les aplicacions

$$\begin{array}{ccc} E_1 & \to & E_1 \oplus E_2 \\ u & \mapsto & (u, \vec{0}) \end{array} \qquad \begin{array}{ccc} E_2 & \to & E_1 \oplus E_2 \\ v & \mapsto & (\vec{0}, v) \end{array}$$

$$\begin{array}{ccc} E_1 \oplus E_2 & \to & E_1 \\ (u,v) & \mapsto & u \end{array} \qquad \begin{array}{ccc} E_1 \oplus E_2 & \to & E_2 \\ (u,v) & \mapsto & v \end{array}$$

són lineals.

3. Sigui F un subespai de l'espai vectorial E. L'aplicació

$$\begin{array}{ccc} E & \to & E/F \\ u & \mapsto & [u] \end{array}$$

que fa correspondre a cada vector la seva classe és lineal.

4. Siguin $a_1, \ldots, a_n$ nombres reals fixos. L'aplicació

$$\begin{array}{ccc} \mathbf{R}^n & \longrightarrow & \mathbf{R}^1 \\ (x^1, \ldots, x^n) & \longmapsto & a_1 x^1 + \ldots + a_n x^n \end{array}$$

és lineal. Més general, si a_i^j són nombres reals fixos, $i = 1, \ldots, n$, $j = 1, \ldots, m$, l'aplicació

$$\begin{array}{ccc} \mathbf{R}^n & \longrightarrow & \mathbf{R}^m \\ (x^1, \ldots, x^n) & \longmapsto & (\sum_{i=1}^{n} a_i^1 x^i, \ldots, \sum_{i=1}^{n} a_i^m x^i) \end{array}$$

és lineal.

Sigui $f : E \to F$ una aplicació lineal:

- Anomenarem *nucli de* f el subespai vectorial de E

$$\operatorname{Nuc} f = \{u \in E \mid f(u) = \vec{0}\}.$$

- Anomenarem *imatge de* f el subespai vectorial de F

$$\operatorname{Im} f = \{v \in F \mid \text{ existeix } u \in E \text{ tal que } v = f(u)\}.$$

Exercici:

Proveu que, efectivament, Nuc f i Im f són subespais vectorials.

Proposició 1.1 *Si l'aplicació $f : E \to F$ és lineal i E és de dimensió finita, aleshores* Nuc f *i* Im f *són de dimensió finita i* $\dim E = \dim \text{Nuc}\, f + \dim \text{Im}\, f$.

DEMOSTRACIÓ: Nuc f és un subespai vectorial de E i, per (IV.3.9), és de dimensió finita. Agafem una base de Nuc f, $v_1, \ldots, v_k$, i completem-la fins a obtenir una base de E (IV.3.8): $v_1, \ldots, v_k, \ldots, v_n$. Les imatges per f dels k primers vectors són $\vec{0}$. Les imatges

$$f(v_{k+1}), \ldots, f(v_n)$$

formen una base de Im f. En efecte, generen Im f, ja que si $v \in \text{Im}\, f$, existeix un $u = \sum_{i=1}^{n} a^i v_i \in E$ tal que

$$v = f(u) = f\left(\sum_{i=1}^{n} a^i v_i\right) = \sum_{i=1}^{n} a^i f(v_i) = \sum_{i=k+1}^{n} a^i f(v_i).$$

A més, són linealment independents, ja que

$$\vec{0} = \sum_{i=k+1}^{n} a^i f(v_i) = f(\sum_{i=k+1}^{n} a^i v_i) \Rightarrow \sum_{i=k+1}^{n} a^i v_i \in \text{Nuc}\, f \Rightarrow$$

$$\Rightarrow \sum_{i=k+1}^{n} a^i v_i = \sum_{i=1}^{k} b^i v_i \Rightarrow \sum_{i=1}^{k} b^i v_i - \sum_{i=k+1}^{n} a^i v_i = \vec{0}.$$

Els vectors que apareixen en aquesta combinació lineal són linealment independents; per tant, els coeficients són 0. En particular $a^i = 0$ per a $i = k+1, \ldots, n$. □

Es diu *rang d'una aplicació lineal* f la dimensió de la seva imatge.

Una aplicació lineal injectiva es diu un *monomorfisme* . Una aplicació lineal exhaustiva es diu un *epimorfisme.* Una aplicació lineal bijectiva es diu un *isomorfisme.* Una aplicació lineal d'un espai E en ell mateix, $E \to E$, es diu un *endomorfisme.* Un endomorfisme bijectiu es diu un *automorfisme.*

Proposició 1.2 *Una aplicació lineal f és injectiva si i només si* Nuc $f = \vec{0}$. *Una aplicació lineal f és exhaustiva si i només si* Im $f = F$.

DEMOSTRACIÓ: La segona afirmació no és res més que la definició d'exhaustivitat. Provem la primera.

$\Longrightarrow$) $u \in \text{Nuc}\, f \Rightarrow f(u) = \vec{0} = f(\vec{0}) \Rightarrow$ (per ser f injectiva) $u = \vec{0}$.

$\Longleftarrow$) $f(u) = f(v) \Rightarrow f(u - v) = \vec{0}$ (per ser f lineal) $\Rightarrow$

$\Rightarrow u - v \in \text{Nuc}\, f \Rightarrow u - v = \vec{0} \Rightarrow u = v$.□

Proposició 1.3 *Si f és lineal i bijectiva, aleshores f^{-1} també és lineal.*

DEMOSTRACIÓ: Per provar que $f^{-1}(u+v) = f^{-1}(u) + f^{-1}(v)$, només cal veure que, per l'aplicació bijectiva f, els dos costats tenen la mateixa imatge,

$$f(f^{-1}(u+v)) = u+v = ff^{-1}(u) + ff^{-1}(v) = f(f^{-1}(u) + f^{-1}(v)).$$

De manera semblant es veu que $f^{-1}(au) = af^{-1}(u)$.□

Observació:

Aquesta proposició ens diu que si $f : E \to F$ és un isomorfisme, hi ha una aplicació lineal $g : F \to E$ ($g = f^{-1}$) tal que $g \circ f = I_E$ i $f \circ g = I_F$. (Aquí, $I_E : E \to E$ és l'aplicació identitat, que envia tot element a ell mateix; anàlogament per a I_F.)

Recíprocament, si existeix una aplicació lineal $g : F \to E$ tal que $g \circ f = I_E$ i $f \circ g = I_F$, resulta que f és injectiva i exhaustiva i, per tant, és bijectiva i inversa de g. Aleshores (1.3) ens diu que f és un isomorfisme.

Dos espais vectorials E i F es diuen *isomorfs* si existeix un isomorfisme $E \to F$. Escriurem llavors

$$E \cong F.$$

Exemples:

5. Si E és un espai vectorial de dimensió n sobre K, aleshores $E \cong K^n$ (exemple 1).
6. Sigui $E_1 \oplus E_2$ la suma directa de dos espais vectorials E_1 i E_2. Considerem els subespais

$$E_1' = \{(u, \vec{0}) \mid u \in E_1\} \qquad E_2' = \{(\vec{0}, v) \mid v \in E_2\}.$$

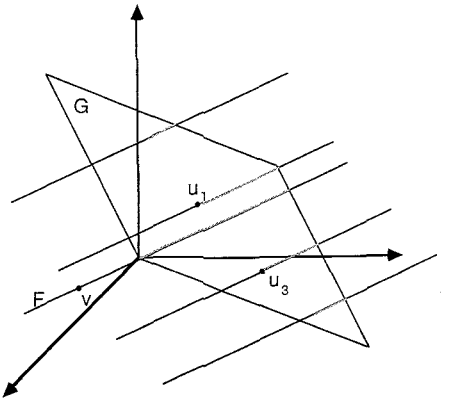
Les aplicacions

$$\begin{array}{ccc} E_1 & \to & E_1' \\ u & \mapsto & (u, \vec{0}) \end{array} \qquad \begin{array}{ccc} E_2 & \to & E_2' \\ v & \mapsto & (\vec{0}, v) \end{array}$$

són isomorfismes. Els subespais E_1' i E_2' tenen intersecció $\{(\vec{0}, \vec{0})\}$ i la seva suma és tot $E_1 \oplus E_2$; és a dir,

$$E_1 \oplus E_2 = E_1' \oplus E_2' \quad \text{amb} \quad E_i' \cong E_i,\ i = 1, 2.$$

Tota suma directa d'espais és també, doncs, suma directa de subespais isomorfs als espais de partença.



7. Siguin F i G dos subespais complementaris a E; és a dir, $F \oplus G = E$. A la figura hem representat el cas particular en què $F = \langle v \rangle$ és una recta i $G = \langle u_1, u_2 \rangle$ un pla. Aleshores, $\mathbf{R}^3/F$ és el conjunt de rectes paral·leles a F (IV.6). Aquest conjunt està en correspondència bijectiva clara amb els punts del pla G: a cada punt de G hi associem la recta paral·lela a F que passa per aquest punt. En general,

$$\begin{array}{ccc} G & \to & E/F \\ u & \mapsto & [u] \end{array}$$

és una aplicació lineal de nucli $G \cap F = \{\vec{0}\}$ i exhaustiva, ja que la classe d'un $w \in E, w = u + v, u \in G, v \in F$, és $[w] = [u]$ i, per tant, és imatge del vector u de G. Els espais G i E/F són, doncs, isomorfs.

Proposició 1.4 *Dos espais vectorials de dimensió finita E i F són isomorfs si i només si* $\dim E = \dim F$.

DEMOSTRACIÓ: Si $E \cong F$, tenen la mateixa dimensió com a conseqüència immediata de (1.2) i (1.1).

Suposem que $\dim E = \dim F$. Escollim bases $u_1, \ldots, u_n$ de E i $v_1, \ldots, v_n$ de F. L'aplicació

$$f : E \to F$$

definida per

$$f\left(\sum_{i=1}^{n} a^i u_i\right) = \sum_{i=1}^{n} a^i v_i$$

és clarament lineal i exhaustiva. Per provar la injectivitat fem servir (1.2) □

Exemple:

$\mathbf{R}^2$ i $\mathbf{C}$ són espais vectorials sobre $\mathbf{R}$ isomorfs. Un isomorfisme és

$$(a, b) \mapsto a + bi.$$

V.2 Matriu associada a una aplicació lineal

Una aplicació lineal queda totalment determinada per les imatges dels vectors d'una base. Aquestes imatges poden ésser, però, qualssevol. Demostrem aquestes afirmacions.

Proposició 2.1 *Sigui $B = \{u_i \mid i \in I\}$ una base d'un espai vectorial E sobre el cos K. Siguin $\{w_i \mid i \in I\}$ vectors qualssevol d'un espai vectorial F sobre K. Existeix una aplicació lineal, i només una,*

$$f : E \to F,$$

tal que $f(u_i) = w_i$ per a cada $i \in I$.

DEMOSTRACIÓ: Com que una aplicació lineal compleix

$$f\left(\sum_{i=1}^{n} a^i u_i\right) = \sum_{i=1}^{n} a^i f(u_i),$$

l'aplicació que busquem ha de definir-se forçosament per

$$f\left(\sum_{i=1}^{n} a^i u_i\right) = \sum_{i=1}^{n} a^i w_i.$$

Només resta demostrar que, així definida, f és lineal. Feu-ho. □

Els següents fets són fàcils de demostrar (notació com a (2.1)):

f és injectiva $\Leftrightarrow$ $\{w_i \mid i \in I\}$ és linealment independent;
f és exhaustiva $\Leftrightarrow$ $\{w_i \mid i \in I\}$ genera F;
f és bijectiva $\Leftrightarrow$ $\{w_i \mid i \in I\}$ és una base de F.

Sigui $f : E \to F$ lineal amb E i F de dimensió finita. Siguin $u_1, \ldots, u_n$ i $v_1, \ldots, v_n$ bases de E i F respectivament. (2.1) ens diu que f queda determinada perfectament si coneixem les coordenades de $f(u_1), \ldots, f(u_n)$:

$$f(u_i) = \sum_{j=1}^{m} a_i^j v_j \qquad i = 1, \ldots, n.$$

Es diu *matriu associada a f respecte a les bases $\{u_i\}, \{v_j\}$*

$$A = \begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \ldots & a_n^m \end{pmatrix}.$$

Observem que (2.1) ens diu, també, que tota matriu $m \times n$ és la matriu associada a una aplicació lineal respecte a les bases $\{u_i\}$ i $\{v_j\}$.

La matriu associada és l'eina amb què, molt sovint, estudiem una aplicació lineal. Vegem, de moment, com aquesta matriu ens permet calcular les coordenades de la imatge d'un vector. Suposem que $w = \sum_{i=1}^{n} w^i u_i \in E$; aleshores,

$$f(w) = \sum_{i=1}^{n} w^i f(u_i) = \sum_{i=1}^{n} w^i \left(\sum_{j=1}^{m} a_i^j v_j \right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_i^j w^i \right) v_j.$$

Les coordenades $(\bar{w}^1, \ldots, \bar{w}^m)$ de $f(w)$ en la base $\{v_j\}$ són, doncs,

$$\bar{w}^j = \sum_{i=1}^{n} a_i^j w^i, \qquad j = 1, \ldots, m.$$

Escrivim les coordenades $(w^1, \ldots, w^n)$ i $(\bar{w}^1, \ldots, \bar{w}^m)$ com a matrius d'una columna W i $\bar{W}$, respectivament. Aleshores, les igualtats anteriors es poden resumir en

$$\bar{W} = AW.$$

Exemples:

1. L'aplicació

$$\begin{array}{rcl} f : \mathbf{R}^n & \longrightarrow & \mathbf{R}^m \\ (x^1, \ldots, x^n) & \longmapsto & (\sum_{i=1}^n a_i^1 x^i, \ldots, \sum_{i=1}^n a_i^m x^i) \end{array}$$

en les bases $\{e_i = (0, \ldots, 1, \ldots, 0)\}$ té per matriu

$$\begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \ldots & a_n^m \end{pmatrix}.$$

2. Sigui F un subespai de E i

$$f : E \to E/F$$

l'aplicació que fa correspondre a cada vector $u \in E$ la seva classe $[u]$ mòdul F. Sigui $v_1, \ldots, v_k$ una base del subespai F i $v_1, \ldots, v_k, \ldots, v_n$ una base de E; sabem que, aleshores, $[v_{k+1}], \ldots, [v_n]$ és una base de E/F. Respecte a aquestes bases, la matriu de f és

$$\begin{pmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

k columnes $\qquad$ $n-k$ columnes.

3. Considerem la matriu de la identitat $I_E : E \to E$. Agafem en el primer espai E una base $u_1, \ldots, u_n$ i en el segon una base $v_1, \ldots, v_n$. La matriu de I_E respecte a aquestes bases està formada pels coeficients de

$$I_E(u_i) = u_i = \sum_{j=1}^{n} a_i^j v_j.$$

Si w té coordenades $(w^1, \ldots, w^n)$ en la base $\{u_i\}$, les coordenades de $w = I_E(w)$ en la base $\{v_i\}$, $(\bar{w}^1, \ldots, \bar{w}^n)$, compleixen

$$\bar{W} = AW.$$

Tornem, així, a trobar l'expressió del canvi de coordenades de (IV.7). La matriu A és la matriu del canvi.

4. La matriu de $I_E : E \to E$ considerant la mateixa base $\{u_i\}$ en els dos espais, és la matriu identitat

$$I = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

5. Sigui $f : E \to F$ un isomorfisme. Considerem una base $\{u_i\}$ de E i la base $\{f(u_i)\}$ de F. La matriu de f respecte a aquestes bases és la matriu identitat.

Proposició 2.2 *Siguin $f : E \to F$ i $g : F \to H$ aplicacions lineals. Siguin $\{u_1, \ldots, u_n\}$, $\{v_1, \ldots, v_m\}$, $\{e_1, \ldots, e_s\}$ bases de E, F, H, respectivament. Siguin A, B, C les matrius de f, g, $g \circ f$ respecte a aquestes bases. Aleshores*

$$C = BA.$$

DEMOSTRACIÓ: Tenim

$$f(u_i) = \sum_{j=1}^{m} a_i^j v_j, \qquad g(v_j) = \sum_{k=1}^{s} b_j^k e_k,$$

$$\begin{aligned} (g \circ f)(u_i) &= g\left(\sum_j a_i^j v_j\right) = \sum_j a_i^j g(v_j) = \\ &= \sum_j a_i^j \left(\sum_k b_j^k e_k\right) = \sum_k \left(\sum_j a_i^j b_j^k\right) e_k. \end{aligned}$$

Per tant,

$$c_i^j = \sum_j a_i^j b_j^k,$$

és a dir,

$$C = BA.\square$$

Corol·lari 2.3 *El producte de matrius és associatiu.*

DEMOSTRACIÓ: Diguem, primer de tot, que hem fet un abús de llenguatge en dir "producte de matrius" com si es tractés d'una operació. Recordem (IV.7) que aquest producte només està definit quan el nombre de columnes de la primera matriu coincideix amb el de files de la segona. Siguin, doncs,

$$A \in M_{m\times n}(K), \qquad B \in M_{s\times m}(K), \qquad C \in M_{t\times s}(K);$$

considerem espais vectorials E, F, H, G sobre K de dimensions n, m, s, t respectivament. Fixem bases a cada un d'aquests espais. Existeixen, aleshores, aplicacions lineals f, g, h amb matrius A, B, C respecte a aquestes bases. El fet que $h \circ (g \circ f) = (h \circ g) \circ f$ implica, per (2.2), que

$$C(BA) = (CB)A.\square$$

Volem ara relacionar les matrius d'una mateixa aplicació lineal respecte a diferents bases. Sigui, doncs,

$$f : E \to F$$

amb matriu A respecte a les bases $u_1, \ldots, u_n$ de E i $v_1, \ldots, v_m$ de F, i amb matriu B respecte a les bases $\bar{u}_1, \ldots, \bar{u}_n$ de E i $\bar{v}_1, \ldots, \bar{v}_m$ de F. Escrivim f com la composició

$$\begin{array}{ccccccc} & & I_E & & f & & I_F \\ E & \to & E & \to & F & \to & F. \end{array}$$

Considerem en cada un d'aquests quatre espais les bases $\{\bar{u}_i\}$, $\{u_i\}$, $\{v_i\}$, $\{\bar{v}_i\}$. La proposició (2.2) ens diu que

$$B = QAP,$$

on P és la matriu de I_E

$$\bar{u}_i = \sum_{k=1}^{n} p_i^k u_k \qquad i = 1, \ldots, n$$

i Q és la matriu de I_F

$$v_j = \sum_{k=1}^{m} q_j^k \bar{v}_k \qquad j = 1, \ldots, m.$$

En moltes ocasions, el que coneixem són les coordenades de la nova base $\bar{v}_k$ en la base v_j,

$$\bar{v}_k = \sum_{j=1}^{m} r_k^j v_j \qquad k = 1, \ldots, m;$$

recordem, però, que vam veure a (IV.7) que les matrius $R = (r_k^j)$ i Q són inverses una de l'altra. Així doncs,

$$B = R^{-1}AP.$$

V.3 Teorema d'isomorfisme

Teorema 3.1 (d'isomorfisme) . *Si $f : E \to F$ és lineal, aleshores*

$$\operatorname{Im} f \cong E/\operatorname{Nuc} f.$$

DEMOSTRACIÓ: L'aplicació f envia tots els elements d'una classe $u + \operatorname{Nuc} f$ al mateix vector de F; en efecte, si $w \in \operatorname{Nuc} f$, $f(w) = \vec{0}$ i

$$f(u + w) = f(u) + f(w) = f(u).$$

Això ens permet definir una aplicació

$$\begin{array}{ccc} E/\operatorname{Nuc} f & \longrightarrow & \operatorname{Im} f \\ [u] & \longmapsto & f(u) \end{array}$$

que és, clarament, lineal i exhaustiva. Aplicarem (1.2) per veure que és injectiva; si $[u]$ va a $\vec{0}$, vol dir que $f(u) = \vec{0}$, és a dir, $u \in \operatorname{Nuc} f$ i $[u] = [\vec{0}]$. □

Corol·lari 3.2 *Si F i G són subespais de E, es compleix*

$$(F + G)/F \cong G/F \cap G.$$

DEMOSTRACIÓ: L'aplicació

$$\begin{array}{ccc} f : G & \longrightarrow & (F + G)/F \\ v & \longmapsto & [v] \end{array}$$

és lineal; el seu nucli està format pels $v \in G$ tals que $[v] = F$, és a dir, tals que $v \in F$; per tant, $\operatorname{Nuc} f = F \cap G$. A més a més, f és exhaustiva, ja que donada $[u]$, podem escriure $u = w + v$ amb $w \in F$, $v \in G$; aleshores, $[u] = [v] = f(v)$. Aplicant (3.1), obtenim

$$G/F \cap G \cong (F + G)/F. \square$$

Corol·lari 3.3 *Si $F \subset G$ són subespais de E, aleshores*

$$(E/F)/(G/F) \cong E/G.$$

DEMOSTRACIÓ: Observem, primer, que $F \subset G$ implica que si una classe $[u]$ de E/F té un representant a G tots els seus elements són de G. Per tant, G/F és un subconjunt de E/F.

Definim, ara,

$$\begin{array}{rccc} f : E/F & \longrightarrow & E/G \\ [u] & \longmapsto & \{u\}, \end{array}$$

on $\{u\}$ indica la classe de u mòdul G. Aquesta aplicació està ben definida, ja que elements equivalents respecte a F també són equivalents respecte a G. f és lineal i exhaustiva i el seu nucli està format per les classes $[u]$ tals que $\{u\} = \{\vec{0}\}$, és a dir, tals que $u \in G$. Així doncs, $\operatorname{Nuc} f = G/F$ i, aplicant (1.3),

$$(E/F)/(G/F) \cong E/G.\square$$

Observació:

Si la dimensió de l'espai E és finita, $\operatorname{Im} f$ i $E/\operatorname{Nuc} F$ tenen la mateixa dimensió ((1.1) i (IV.6.1)). Són, per tant, espais isomorfs (per 1.4). Què ens diu de nou, doncs, la proposició 3.1? En primer lloc, ens diu que el resultat és vàlid per a qualsevol espai vectorial, de dimensió finita o no. Encara més important és, però, el fet que es defineix un isomorfisme d'una manera molt natural i sense intervenció de bases. Una aplicació lineal definida sense fer servir cap base, o que permet una definició d'aquest tipus, es diu una aplicació *canònica*. Tots els isomorfismes establerts en les tres proposicions d'aquest apartat són isomorfismes canònics.

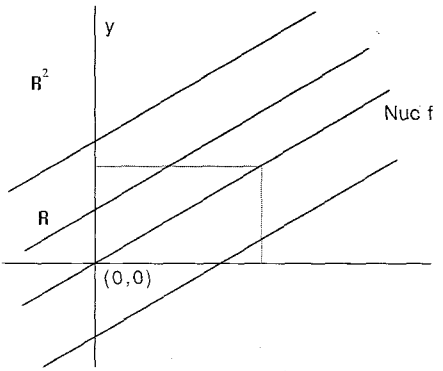
Volem donar, abans d' acabar aquest apartat, uns exemples que il·lustrin el contingut d'aquestes tres proposicions.

Exemples:

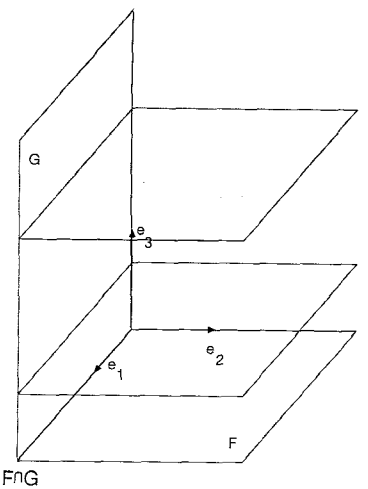
1. Considerem l'aplicació lineal

$$\begin{array}{rccc} f : & \mathbf{R}^2 & \longrightarrow & \mathbf{R}^1 \\ & (x, y) & \longmapsto & x - y. \end{array}$$

El nucli de f correspon en el pla a una recta que passa per $(0,0)$; $\mathbf{R}^2/\operatorname{Nuc} f$ és, aleshores, el conjunt de rectes paral·leles a $\operatorname{Nuc} f$. Cada una d'aquestes rectes talla l'eix de les x en un punt $(a, 0)$, i la seva imatge per f és, precisament, a. L'isomorfisme de (3.1) fa correspondre a cada recta de $\mathbf{R}^2/\operatorname{Nuc} f$ la seva intersecció amb l'eix de les x.



2. Considerem $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$ a $\mathbf{R}^3$. Siguin F i G els subespais vectorials de $\mathbf{R}^3$ definits per $F = \langle e_1, e_2 \rangle$, $G = \langle e_1, e_3 \rangle$. Aleshores $F \cap G = \langle e_1 \rangle$ i $F + G = \mathbf{R}^3$. Per tant, (F+G)/F és el conjunt de plans de $\mathbf{R}^3$ paral·lels a F i $G/F \cap G$ és el conjunt de rectes del pla G paral·leles a $F \cap G$. L'isomorfisme de (3.2) fa correspondre a cada recta de G paral·lela a $F \cap G$, el pla de $\mathbf{R}^3$ paral·lel a F que la conté.



3. Siguin e_1, e_2, e_3 com a l'exemple anterior. Considerem els subespais de $\mathbf{R}^3$, $F = \langle e_1 \rangle$ i $G = \langle e_1, e_2 \rangle$. Aleshores, $\mathbf{R}^3/F$ és el conjunt de rectes de $\mathbf{R}^3$ paral·leles a F i G/F és el conjunt de rectes de G paral·leles

a F. Cada classe de $(\mathbf{R}^3/F)/(G/F)$ està formada per totes les rectes (paral·leles a F) situades en un pla paral·lel a G. L'isomorfisme de (3.3) fa correspondre a cada una d'aquestes classes el pla paral·lel a G que la conté (que és un element de $\mathbf{R}^3/G$).

V.4 L'espai de les aplicacions lineals

Considerem el conjunt $L(E,F)$ de totes les aplicacions lineals de E a F. Hi ha una manera natural de definir una suma i un producte per elements del cos K a $L(E,F)$. Concretament, si $f,g \in L(E,F)$ i $a \in K$ definim *la suma* $f+g$ per

$$(f+g)(u) = f(u) + g(u) \qquad \forall u \in E$$

i *el producte* af per

$$(af)(u) = af(u) \qquad \forall u \in E.$$

Les aplicacions $f+g$ i af són, clarament, lineals. $L(E,F)$ amb aquestes dues operacions compleix totes les condicions d'espai vectorial (IV.1); en direm *espai vectorial de les aplicacions de* E *a* F.

Proposició 4.1 *Si* E *i* F *són de dimensió finita,* $L(E,F)$ *també i* $\dim L(E,F) = \dim E \cdot \dim F$.

DEMOSTRACIÓ: Sigui $u_1,\ldots,u_n$ una base de E i $v_1,\ldots,v_m$ una base de F. Definim

$$f_{ij} : E \to F \qquad i = 1,\ldots,n,\ j = 1,\ldots,m$$

per $f_{ij}(u_k) = \vec{0}$ si $k \neq i$, $f_{ij}(u_i) = v_j$.

La proposició quedarà demostrada si provem que aquestes $n \cdot m$ aplicacions formen una base de $L(E,F)$. Per a això hem de veure que

- $\{f_{ij}\}$ genera $L(E,F)$. En efecte, sigui $f : E \to F$ lineal; suposem que $f(u_k) = \sum_{j=1}^{m} a_k^j v_j$. Aleshores $f = \sum_{i,j} a_i^j f_{ij}$, ja que $\forall u_k$

$$\left(\sum_{i,j} a_i^j f_{ij}\right)(u_k) = \sum_{i,j} a_i^j f_{ij}(u_k) = \sum_j a_k^j v_j = f(u_k).$$

- $\{f_{ij}\}$ són linealment independents. En efecte,

$$\sum_{i,j} a_i^j f_{ij} = 0 \Rightarrow \sum_{i,j} a_i^j f_{ij}(u_k) = \vec{0} \text{ per a tot } k = 1,\ldots,n.$$

Amb un càlcul com el fet més amunt, obtenim

$$\sum_j a_k^j v_j = \vec{0} \qquad k = 1, \ldots, n$$

i, per tant, $a_k^j = 0$ per a tot $k = 1, \ldots, n$ i tot $j = 1, \ldots, m$. □

Observem que la matriu de f_{ij} respecte a les bases $u_1, \ldots, u_n$ i $v_1, \ldots, v_m$ és la matriu E_i^j formada per 0 a tot arreu, menys a la columna i fila j, on apareix un 1. A (IV.3) vam veure que aquestes matrius formen una base de $M_{m\times n}(K)$. L'aplicació

$$\begin{array}{ccccc} L(E,F) & \longrightarrow & M_{m\times n}(K) & & \\ f_{ij} & \longmapsto & E_i^j & i = 1, \ldots, n & , \quad j = 1, \ldots, m \end{array}$$

és un isomorfisme entre aquests dos espais vectorials. De la demostració de (4.1) es dedueix que aquest isomorfisme associa a cada aplicació lineal f la seva matriu en les bases considerades.

V.5 L'àlgebra d'endomorfismes

Un cas particular de l'espai estudiat a l'apartat anterior és $L(E,E)$, l'espai dels endomorfismes de E, que denotarem per $\mathrm{End}(E)$.

Dos elements $f, g \in \mathrm{End}(E)$ es poden compondre sempre i la composició $g \circ f$ és també un element de $\mathrm{End}(E)$ que anomenarem *producte,* o *producte intern* si hi ha perill de confusió amb el producte per elements del cos. Aquest producte compleix les propietats següents:

- Associativa: $h \circ (g \circ f) = (h \circ g) \circ f \qquad \forall f, g, h \in \mathrm{End}(E)$.
- Hi ha un element neutre, que és l'aplicació identitat I_E:

$$I_E \circ f = f \circ I_E = f \qquad \forall f \in \mathrm{End}(E).$$

En general, però, el producte no és commutatiu i els únics elements que tenen invers són els endomorfismes bijectius (automorfismes).

El producte intern de $\mathrm{End}(E)$ està relacionat amb les dues operacions de l'estructura vectorial per les propietats següents:

- Distributives:

$$\begin{array}{rcl} (h+g)\circ f & = & h\circ f + g\circ f \\ h\circ(g+f) & = & h\circ g + h\circ f \end{array} \qquad \forall f, g, h \in \mathrm{End}(E).$$

- $(ag) \circ f = a(g \circ f) = g \circ (af) \quad \forall a \in K \quad \forall f, g \in \mathrm{End}(E)$.

En particular, $\mathrm{End}(E)$ amb les operacions $+$ i $\circ$ és un anell amb unitat.

Un conjunt A amb tres operacions - una suma $+$, un producte $\cdot$, i un producte per elements d'un cos K- es diu una *àlgebra sobre* K si A amb $+$ i $\cdot$ és un anell, A amb $+$ i el producte per elements de K és un espai vectorial i $k(a \cdot b) = (ka) \cdot b = a \cdot (kb) \qquad \forall k \in K, \forall a, b \in A$.

Exemples:

1. El conjunt de polinomis $K[x]$ és una àlgebra commutativa i amb unitat sobre K.
2. $\mathrm{End}(E)$ és una àlgebra amb unitat. També és una àlgebra amb unitat el conjunt de les matrius quadrades $M_{n\times n}(K)$. Si E és de dimensió n, a l'apartat anterior hem establert una aplicació

$$\begin{array}{rccc} \mathrm{End}(E) = & L(E,E) & \longrightarrow & M_{n\times n}(K) \\ & f & \longmapsto & A \end{array}$$

on A és la matriu associada a f en una base prefixada de E. Aquesta aplicació és un isomorfisme d'espais vectorials i, a més, "conserva" els productes interns per (2.2). Es diu, aleshores, que és un isomorfisme d'àlgebres i que les àlgebres $\mathrm{End}(E)$ i $M_{n\times n}(K)$ són isomorfes.

Anomenarem *homotècia vectorial de raó* a l'endomorfisme aI_E de E.

Proposició 5.1 *Si E té dimensió 1, els seus únics endomorfismes són les homotècies vectorials.*

DEMOSTRACIÓ: Sigui $u \neq 0$ una base de E i $f \in \mathrm{End}(E)$. La imatge de u s'expressarà com a combinació lineal de la base

$$f(u) = au.$$

Aleshores, la imatge de qualsevol altre vector $v = cu \in E$ és

$$f(v) = f(cu) = cf(u) = c(au) = a(cu) = av,$$

d'on resulta que $f = aI_E$. □

Observem que en el cas de la proposició 5.1 la matriu de $f = aI_E$ és precisament (a) i, per tant, independent de la base. L'isomorfisme d'àlgebres de l'exemple 2 és, en aquest cas, un isomorfisme canònic

$$\begin{array}{ccc} \mathrm{End}(E) & \xrightarrow{\simeq} & K \\ f = aI_E & \longmapsto & a. \end{array}$$

V.6 L'espai dual

En aquest apartat anem a estudiar un altre cas particular de l'espai d'aplicacions lineals: el cas en què el segon espai vectorial és K. Les aplicacions lineals a K es diuen, també, *formes*; de l'espai

$$E' = L(E, K)$$

en direm l'*espai dual de* E . Totes les consideracions de l'apartat 4 s'apliquen, en particular, a aquest cas. Així doncs, E' és un espai vectorial de la mateixa dimensió que E (si $\dim E$ és finita). Donada una base $u_1, \ldots, u_n$ de E, les aplicacions

$$\begin{array}{rccccl} u_i' : & E & \longrightarrow & K & & \\ & u_j & \longmapsto & 0 & si \quad i \neq j & \\ & u_j & \longmapsto & 1 & si \quad i = j & \qquad i = 1, \ldots, n \end{array}$$

formen una base de E', que anomenarem *base dual* de $u_1, \ldots, u_n$.

Atenció!:

Suposem que $u_1, \ldots, u_n$ i $v_1, \ldots, v_n$ són dues bases diferents de E, però amb alguns vectors comuns, per exemple $u_1 = v_1$; en les bases duals $u_1', \ldots, u_n'$, $v_1', \ldots, v_n'$ els elements u_1' i v_1' no tenen per què ésser iguals.

Proposició 6.1 *Sigui $u_1, \ldots, u_n$ una base de E i $u_1', \ldots, u_n'$ la seva base dual. Les coordenades d'una forma $\omega \in E'$ en la base $u_1', \ldots, u_n'$ són $\omega(u_1), \ldots, \omega(u_n)$*

$$\omega = \omega(u_1)u_1' + \ldots + \omega(u_n)u_n'.$$

DEMOSTRACIÓ: Per a tot vector u_k de la base de E, tenim

$$\left(\sum_{i=1}^{n} \omega(u_i)u_i'\right)(u_k) = \sum_{i=1}^{n} \omega(u_i)u_i'(u_k) = \omega(u_k)$$

i, per tant, $\displaystyle\sum_{i=1}^{n} \omega(u_i)u_i' = \omega$. □

Fixada una aplicació lineal $f : E \to F$, cada element $\omega \in F'$ ens dóna, en compondre amb f, un element $\omega \circ f$ de E'

$$\begin{array}{ccc} & f & \\ E & \longrightarrow & F \\ & \searrow & \downarrow \omega \\ & \omega \circ f & K. \end{array}$$

Tenim, per tant, una aplicació de F' en E', que designarem per

$$f' : F' \to E'$$

DEMOSTRACIÓ: φ és lineal, ja que per a tot $u, v \in E$ i tot $\omega \in E'$,

$$\begin{aligned}(\varphi(u+v))(\omega) &= \langle \omega, u+v\rangle = \omega(u+v) = \\ &= \omega(u)+\omega(v) = \langle \omega, u\rangle + \langle \omega, v\rangle = \\ &= \varphi(u)(\omega)+\varphi(v)(\omega) = \\ &= (\varphi(u)+\varphi(v))(\omega),\end{aligned}$$

d'on s'obté que $\varphi(u+v) = \varphi(u)+\varphi(v)$.

Anàlogament es prova que $\varphi(au) = a\varphi(u)$.

Per veure que φ és injectiva, provarem que l'únic vector del nucli és $\vec{0}$ (1.2). Si $\varphi(u) = \vec{0}$, aleshores, per a tot $\omega \in E'$,

$$0 = \varphi(u)(\omega) = \langle \omega, u\rangle = \omega(u).$$

Si $u \neq \vec{0}$, hi ha una base de E de la forma $u, u_2, \ldots, u_n$; considerem un $\omega \in E'$ tal que $\omega(u) = 1, \omega(u_i) = 0$, $i = 2, \ldots, n$. Aleshores

$$\varphi(u)(\omega) = \langle \omega, u\rangle = 1 \neq 0,$$

en contra del que hem obtingut abans. Així doncs, $u = \vec{0}$ i l'aplicació φ és injectiva.

L'exhaustivitat de φ resulta de (1.1) i del fet que E i E'' tenen la mateixa dimensió finita. □

Observació:

En la demostració anterior només hem fet servir que la dimensió de E és finita per provar l'exhaustivitat; per a espais de dimensió infinita, φ és un monomorfisme.

Proposició 6.4 *Sigui $f : E \to F$ una aplicació lineal entre espais de dimensió finita, i sigui $f'' : E'' \to F''$ la seva bidual. Si $\varphi : E \cong E''$ i $\bar{\varphi} : F \cong F''$ són els isomorfismes de (6.3), aleshores*

$$\bar{\varphi}^{-1} \circ f'' \circ \varphi : E \cong E'' \longrightarrow F'' \cong F$$

coincideix amb f.

DEMOSTRACIÓ: Si $u \in E$,

$$(f'' \circ \varphi)(u) = f''(\langle\ , u\rangle) = \langle\ , u\rangle \circ f' \in F'',$$

i si $\omega \in F'$,

$$\begin{aligned}(\langle\ , u\rangle \circ f')(\omega) &= \langle\ , u\rangle(f'(\omega)) = \langle f'(\omega), u\rangle = (f'(\omega))(u) = \\ &= \omega(f(u)) = \langle \omega, f(u)\rangle = \langle\ , f(u)\rangle(\omega),\end{aligned}$$

d'on resulta que

$$\langle\ , u\rangle \circ f' = \langle\ , f(u)\rangle.$$

El vector de F que correspon per $\bar{\varphi}$ a aquest element de F'' és $f(u)$ i, per tant,

$$\left(\bar{\varphi}^{-1} \circ f'' \circ \varphi\right)(u) = \bar{\varphi}^{-1}\left(\langle\ , f(u)\rangle\right) = f(u).$$

Així, doncs, $\bar{\varphi}^{-1} \circ f'' \circ \varphi = f$. □

V.7 Subespais ortogonals

En tot aquest apartat suposarem que treballem només amb espais vectorials de dimensió finita. Com a l'apartat anterior, E' indicarà el dual de l'espai E, $E' = L(E, K)$.

Sigui A un subconjunt de E. Definim l'*ortogonal de* A com el conjunt

$$A^{\perp} = \{\omega \in E' \mid \omega(u) = 0\ \forall u \in A\}.$$

Es tenen, aleshores, les propietats següents:

1. $A^{\perp}$ és un subespai vectorial de E'.
2. $A \subset B \Rightarrow B^{\perp} \subset A^{\perp}$.
3. Si F és un subespai de E, $\dim F^{\perp} = \dim E - \dim F$.
4. $E^{\perp} = \{0\}$, $\{\vec{0}\}^{\perp} = E'$.

1, 2 i 4 són immediates; demostrem 3: agafem una base $u_1, \ldots, u_k$ de F i completem-la fins a obtenir una base $u_1, \ldots, u_k, u_{k+1}, \ldots, u_n$ de E.

Sigui $u'_1, \ldots, u'_k, u'_{k+1}, \ldots, u'_n$ la base dual corresponent. Per a $j = k+1, \ldots, n$, u'_j s'anul·la sobre la base $u_1, \ldots, u_k$ de F i, per tant, sobre tot F; és a dir, $u'_j \in F^{\perp}$ per a $j = k+1, \ldots, n$. Ara bé, aquests elements formen una base de $F^{\perp}$, ja que són linealment independents (en ésser part d'una base) i generen $F^{\perp}$, ja que si $\omega = a^1 u'_1 + \ldots + a^n u'_n \in F^{\perp} \subset E'$, com que $u_h \in F$, $h = 1, \ldots, k$, es té $0 = \omega(u_h) = a^h$, d'on $\omega = a^{k+1} u'_{k+1} + \ldots + a^n u'_n$.

Volem definir, ara, l'ortogonal d'un subconjunt A de E'. Hi ha dues maneres de fer-ho.

1. $A^{\perp} = \{u \in E \mid \omega(u) = 0\ \forall \omega \in A\}$.

   Obtenim, així, un subconjunt de l'espai inicial E d'una manera molt semblant a com hem obtingut l'ortogonal d'un $A \subset E$.

2. Apliquem la definició que hem donat al començament de l'apartat. Obtenim, així, un subconjunt del bidual

$$A^{\perp} = \{\alpha \in E'' \mid \alpha(\omega) = 0\ \forall \omega \in A\}.$$

Aquestes dues maneres són, essencialment, la mateixa. Amb més precisió, aquests dos ortogonals es corresponen per l'isomorfisme de (6.3). En efecte, recordem que en aquell isomorfisme un element $\alpha \in E''$ corresponia a un vector $u \in E$ de forma que $\alpha = \langle\ ,u\rangle$. Per tant, a 2,

$$\begin{aligned}\{\alpha \in E'' \mid \alpha(\omega) = 0\ \forall \omega \in A\} &= \{\langle\ ,u\rangle \in E'' \mid \langle \omega, u\rangle = 0\ \forall \omega \in A\} = \\ &= \{\langle\ ,u\rangle \in E'' \mid \omega(u) = 0\ \forall \omega \in A\},\end{aligned}$$

que correspon a $\{u \in E \mid \omega(u) = 0\ \forall \omega \in A\}$ de 1.

La definició 2 és la donada al començament de l'apartat. Per tant, les propietats 1, 2, 3 i 4 són vàlides en aquest cas. Les consideracions anteriors ens diuen que aquestes propietats són també vàlides per a ortogonals definits segons 1. En endavant treballarem sempre amb la definició 1.

Altres propietats dels ortogonals són:

5. Si F és un subespai vectorial, $F^{\perp\perp} = F$.

6. Si F i G són subespais vectorials,

$$(F \cap G)^{\perp} = F^{\perp} + G^{\perp} \text{ i } (F + G)^{\perp} = F^{\perp} \cap G^{\perp}.$$

7. Si $E = F \oplus G$, $E' = F^{\perp} \oplus G^{\perp}$.

DEMOSTRACIÓ: de 5.

$$u \in F \Rightarrow \langle \omega, u\rangle = \omega(u) = 0\ \forall \omega \in F^{\perp} \Rightarrow u \in F^{\perp\perp},$$

d'on $F \subset F^{\perp\perp}$; però la propietat 3 ens diu que aquests dos espais tenen la mateixa dimensió i, per tant, $F = F^{\perp\perp}$. □

DEMOSTRACIÓ: de 6.

$$\begin{aligned} F \cap G \subset F,\ F \cap G \subset G &\Rightarrow \text{(per 2) } F^{\perp} \subset (F \cap G)^{\perp}, G^{\perp} \subset (F \cap G)^{\perp} \Rightarrow \\ &\Rightarrow F^{\perp} + G^{\perp} \subset (F \cap G)^{\perp}; \qquad (*) \\ F + G \supset F, F + G \supset G &\Rightarrow \text{(per 2) } F^{\perp} \supset (F + G)^{\perp}, G^{\perp} \supset (F + G)^{\perp} \Rightarrow \\ &\Rightarrow (F + G)^{\perp} \subset F^{\perp} \cap G^{\perp}; \qquad (**) \end{aligned}$$

aleshores, per 5, $(*)$ i $(**)$,

$$F \cap G = (F \cap G)^{\perp\perp} \subset (F^{\perp} + G^{\perp})^{\perp} \subset F^{\perp\perp} \cap G^{\perp\perp} = F \cap G,$$

i totes les inclusions són igualtats; en particular, $F \cap G = (F^{\perp} + G^{\perp})^{\perp}$, d'on, per 5, $(F \cap G)^{\perp} = F^{\perp} + G^{\perp}$.

De manera semblant es procedeix per demostrar l'altra igualtat. □

DEMOSTRACIÓ: de 7.

$$\begin{aligned} E = F \oplus G &\Leftrightarrow E = F + G \text{ i } F \cap G = \{\vec{0}\} \Leftrightarrow \\ &\Leftrightarrow \{\vec{0}\} = E^{\perp} = (F+G)^{\perp} = F^{\perp} \cap G^{\perp} \text{ i} \\ &\quad E' = \{\vec{0}\}^{\perp} = (F \cap G)^{\perp} = F^{\perp} + G^{\perp} \\ &\Leftrightarrow E' = F^{\perp} \oplus G^{\perp}. \square \end{aligned}$$

Proposició 7.1 *Sigui $f : E \to F$ una aplicació lineal entre espais vectorials de dimensió finita i $f' : F' \to E'$ la seva dual. Llavors,*

$$(\operatorname{Im} f)^{\perp} = \operatorname{Nuc} f' \qquad i \ (\operatorname{Nuc} f)^{\perp} = \operatorname{Im} f'.$$

DEMOSTRACIÓ:

$$\begin{aligned} (\operatorname{Im} f)^{\perp} &= \{\omega \in F' \mid \omega(v) = 0 \ \forall v \in Imf\} = \{\omega \in F' \mid \omega(fu) = 0 \ \forall u \in E\} = \\ &= \{\omega \in F' \mid (f'\omega)(u) = 0 \ \forall u \in E\} = \{\omega \in F' \mid f'\omega = 0\} = \operatorname{Nuc} f'. \\ (Imf')^{\perp} &= \{u \in E \mid \rho(u) = 0 \ \forall \rho \in Imf'\} = \{u \in E \mid (f'\omega)(u) = 0 \ \forall \omega \in F'\} = \\ &= \{u \in E \mid \omega(fu) = 0 \ \forall \omega \in F'\} = \\ &\quad \text{(per un raonament fet a (6.3))} \\ &= \{u \in E \mid (fu) = 0\} = \operatorname{Nuc} f. \end{aligned}$$

La propietat 5 ens dóna, ara, la segona igualtat. □

V.8 Nota històrica

Per Leonhard Euler (1707–1783) una funció era una fórmula o equació que contenia variables i constants. Euler i Joseph-Louis Lagrange (1736–1813) ja coneixien que les solucions d'un sistema homogeni formen un espai vectorial, però això no és explotat fins a Augustin-Louis Cauchy (1789–1857). La teoria d'aplicacions lineals es desenvolupa a mitjans del segle 19 (encara que una definició precisa com l'actual fou donada per Giuseppe Peano (1858–1932) a finals de segle) i la connexió entre matrius i aplicacions lineals fou establerta i desenvolupada per Arthur Cayley (1821–1895) el 1855. Georg Ferdinand Frobenius (1849–1917) considera el 1879 el rang d'una matriu i la utilitza en l'estudi dels sistemes d'equacions lineals. (Veure Cap.VII).

Podeu observar en les notes històriques d'aquests capítols que, contràriament al que hem fet en el llibre, en el desenvolupament històric de les matemàtiques les definicions precises arriben després de la utilització de les eines i de l'obtenció de bona part dels resultats.

V.9 Exercicis

1. Proveu que donada qualsevol aplicació lineal $f : E \longrightarrow F$ existeixen bases de E i F tals que la matriu de f en aquestes bases és

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Quin significat té r?

2. Sigui

$$E = \{\begin{pmatrix} 0 & a \\ b & c \end{pmatrix} \in M_{2\times 2}(\mathbf{R}) \mid c = a + b\}.$$

Considerem l'endomorfisme $f : E \longrightarrow E$ donat per

$$f\begin{pmatrix} 0 & a \\ b & c \end{pmatrix} = \begin{pmatrix} 0 & 3c + 3a \\ -2a - b & a - b + 3c \end{pmatrix}.$$

Trobeu una base de E, la matriu de f en aquesta base i sengles bases de $\operatorname{Nuc} f$ i $\operatorname{Im} f$.

3. Quina és la matriu del canvi de base entre

$$\{1, x, x^2, \ldots, x^n\} \quad \text{i} \quad \{1, x - a, (x - a)^2, \ldots, (x - a)^n\}$$

com a bases de l'espai vectorial dels polinomis reals de grau n? Utilitzeu-la per provar la fórmula de Taylor:

$$p(x) = p(a) + p'(a)(x - a) + \frac{1}{2!}p''(a)(x - a)^2 + \cdots + \frac{1}{n!}p^{(n)}(a)(x - a)^n.$$

4. Sigui $f : E \longrightarrow F$ una aplicació lineal. Demostreu

a) f és injectiva $\Longleftrightarrow$ existeix una $g : F \longrightarrow E$ lineal tal que $g \circ f = I$.

b) f és exhaustiva $\Longleftrightarrow$ existeix una $g : F \longrightarrow E$ lineal tal que $f \circ g = I$.

5. Donades dues aplicacions lineals $f : E \longrightarrow F$ i $g : E \longrightarrow G$, trobeu sengles condicions necessàries i suficients perquè

a) existeixi $h : F \longrightarrow G$ lineal tal que $h \circ f = g$.

b) existeixi una única $h : F \longrightarrow G$ lineal tal que $h \circ f = g$.

c) existeixi un monomorfisme $h : F \longrightarrow G$ tal que $h \circ f = g$.

6. Donat un endomorfisme f d'un espai vectorial de dimensió finita n, demostreu que els conjunts

$$\mathcal{F} = \{g \in \operatorname{End}(E) \mid f \circ g = 0\} \quad \text{i} \quad \mathcal{G} = \{g \in \operatorname{End}(E) \mid g \circ f = 0\}$$

són subepais vectorials $\operatorname{End}(E)$ i determineu-ne les dimensions.

7. Demostreu que tot endomorfisme f d'un espai vectorial, E, de dimensió finita pot expressar-se com diferència de dos automorfismes.

8. Un endomorfisme $f : E \longrightarrow E$ es diu un *projector* si $f^2 = f$. Demostreu:

a) f és un projector si i només si $I - f$ ho és.

b) Si f és un projector, $E = \operatorname{Nuc} f \oplus \operatorname{Im} f$.

c) Si f i g són projectors determineu condicions necessàries i suficients perquè $f + g$ també ho sigui.

d) Si f és un projector, trobeu les relacions existents entre $\operatorname{Nuc} f$, $\operatorname{Im} f$, $\operatorname{Nuc}(I - f)$, $\operatorname{Im}(I - f)$.

9. Sigui $e_1, \ldots, e_n$ una base de l'espai vectorial E i siguin $(a_i^1, \ldots, a_i^n)$ les coordenades dels vectors $v_i \in E$, $i = 1, \ldots, k$. Quines condicions han de cumplir les coordenades d'una forma $\omega \in E'$ en la base dual, $e'_1, \ldots, e'_n$, de $e_1, \ldots, e_n$, perquè $\omega \in \langle v_1, \ldots, v_k \rangle^\perp$?

Siguin ara $(b_j^1, \ldots, b_j^n)$ coordenades de formes $\omega_j \in E'$, $j = 1, \ldots, k$, en la base $e'_1, \ldots, e'_n$. Quines condicions han de cumplir les coordenades d'un vector $v \in E$ perquè $v \in \langle \omega_1, \ldots, \omega_k \rangle^\perp$?

10. Demostreu que $\omega_1, \ldots, \omega_m \in E'$ són linealment independents si i només si per a cada m-pla $(a_1, \ldots, a_m) \in K^m$ existeix un vector $u \in E$ tal que $\omega_i(u) = a_i$, $i = 1, \ldots, m$.

11. Sigui $f \in \operatorname{End}(E)$ tal que $f^2 = I$. Siguin $E_1 = \{x \in E \mid f(x) = x\}$, $E_2 = \{x \in E \mid f(x) = -x\}$. Demostreu que $E = E_1 \oplus E_2$. Vol dir això que per a tot $x \in E$ $f(x) = x$ o $f(x) = -x$?

12. Sigui $f \in \operatorname{End}(E)$. Demostreu que $\operatorname{Nuc} f = \operatorname{Im} f$ si i només si $\dim E$ és parell, $f^2 = 0$ i $\operatorname{rang} f = n/2$.

13. Demostreu que $f \in \operatorname{End}(E)$ commuta amb tots els endomorfismes de E si i només si $f = aI$, $a \in K$.

14. Sigui $f \in \operatorname{End}(E)$ tal que $f^2 + f + I = 0$. Demostreu que f és isomorfisme i determineu el seu invers.

15. Sigui E un espai vectorial de dimensió finita i f un subespai vectorial de E. La inclusió $F^\perp \subset E'$ permet definir una aplicació

$$f : E'' \longrightarrow (F^\perp)'$$

per restricció. Demostreu que f és un epimorfisme.

Sigui ara $\varphi : E \cong E'$ l'isomorfisme de la proposició (6.3). Comproveu que $\operatorname{Nuc}(f \circ \varphi) = F$ i expliciteu un isomorfisme $E/F \cong (F^\perp)'$.

V.10 Exercicis de programar

Els exercicis que proposem a continuació utilitzen tots, en algun moment, l'exercici VII.12.

16. Elaboreu un programa que permeti

a) canviar de base les coordenades d'un vector donat.

b) canviar de base la matriu d'una aplicació lineal donada.

(Indicació: si $u_1, \ldots, u_n$ és la base original i $e_1, \ldots, e_n$ la nova base, construïu la matriu del canvi i la seva inversa —utilitzant l'exercici VII.12—. Utilitzeu la subrutina de l'exercici IV.15c).

17. Elaboreu un programa que, donats $e_1, \ldots, e_n$ vectors de $\mathbf{R}^n$, comprovi que formen una base i en calculi la seva base dual. (Indicació: si P és la matriu que expressa la base $e_1, \ldots, e_n$ en funció de la base canònica, $(P^t)^{-1}$ expressa les formes $e'_1, \ldots, e'_n$ en funció de la base dual de la canònica).

Capítol VI

Determinants

VI.1 Determinant de n vectors

Considerem el conjunt de les matrius $n \times n$ sobre K. Volem associar a cada matriu un element del cos K, el seu "determinant", de forma que es compleixin les següents propietats: si multipliquem per $a \in K$ els elements d'una columna, el determinant queda multiplicat per a; si una columna és suma de dos, el determinant és suma dels determinants calculats amb cada una de les columnes-sumands; si dues columnes són iguals, el determinant és zero. Aquestes tres condicions són, de fet, prou restrictives per no permetre gaire marge a l'hora d'escollir (definir!) qui serà el determinant d'una matriu. Anem a veure-ho.

Observem, primer, que les condicions imposades es refereixen a les columnes; per a això convé considerar cada matriu $n \times n$ com un element de $\overbrace{K^n \times \cdots \times K^n}^{n}$, interpretant cada columna com una n-pla de K^n. Aleshores, un determinant ha d'ésser una aplicació

$$\begin{array}{rccc} \det : & \overbrace{K^n \times \cdots \times K^n}^{n} & \longrightarrow & K \\ & (a_1, \ldots, a_n) & \longmapsto & \det(a_1, \ldots, a_n) \end{array}$$

que compleixi:

- $\det(a_1, \ldots, aa_i, \ldots, a_n) = a \det(a_1, \ldots, a_n) \qquad \forall a \in K, \quad i = 1, \ldots, n.$
- $\det(a_1, \ldots, a_i + a_i', \ldots, a_n) = \det(a_1, \ldots, a_i, \ldots, a_n) + \det(a_1, \ldots, a_i', \ldots, a_n)$ $i = 1, \ldots, n.$
- $\det(a_1, \ldots, a_n) = 0 \qquad$ si $a_i = a_j$ amb $i \neq j$.

Això ens porta a la primera definició que anem a donar per a una situació una mica més general.

Sigui E un espai vectorial de dimensió n sobre un cos K. Una *n-forma lineal alternada* és una aplicació

$$D : \overbrace{E \times \cdots \times E}^{n} \longrightarrow K$$

que compleix

(a) $D(v_1, \ldots, av_i, \ldots, v_n) = a\, D(v_1, \ldots, v_i, \ldots, v_n) \quad \forall a \in K, \qquad i = 1, \ldots, n.$

(b) $D(v_1, \ldots, v_i + v_i', \ldots, v_n) = D(v_1, \ldots, v_i, \ldots, v_n) + D(v_1, \ldots, v_i', \ldots, v_n)$
$i = 1, \ldots, n.$

(c) $D(v_1, \ldots, v_n) = 0$ si $v_i = v_j$ amb $i \neq j$.

Les condicions (a) i (b) es resumeixen dient que D és *multilineal* o bé lineal a cada factor; el nom d'"alternada" es refereix a la tercera condició o, més exactament, a la primera de les propietats que enunciarem. El nom de "forma" es reserva per a aplicacions lineals o multilineals a K.

Passem a enunciar les propietats de les n-formes lineals alternades.

1. $D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$.

   En efecte, per (c),

$$\begin{aligned} 0 &= D(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_n) = \quad \text{(per (b))} \\ &= D(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) + D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) + \\ &\quad + D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) + D(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_n) = \\ &\quad \text{(per (c))} \\ &= D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) + D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n). \square \end{aligned}$$

   Aquesta propietat 1 equival en molts casos a la condició (c). Concretament, si una aplicació $D : E \times \cdots \times E \longrightarrow K$ compleix (a), (b) i 1, aleshores, per a $v_i = v_j$

$$D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$$

   si i només si

$$2\, D(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) = 0.$$

   Sempre que a K $2 \neq 0$, això equival a

$$D(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) = 0.$$

2. Per a tota permutació $\sigma \in \mathcal{S}_n$ de signe $\varepsilon(\sigma)$ (III.2),

$$D(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \varepsilon(\sigma)\, D(v_1, \ldots, v_n).$$

   Aquesta propietat es redueix a 1 si σ és una transposició. En general, podem descompondre σ en producte de transposicions i aplicar 1 reiteradament. $\square$

3. Si un vector $v_i = \vec{0}$, aleshores $D(v_1, \ldots, v_i, \ldots, v_n) = 0$.

   En efecte, com que $v_i = 0v_i$, tenim

$$D(v_1, \ldots, v_i, \ldots, v_n) = 0 \cdot D(v_1, \ldots, v_i, \ldots, v_n) = 0. \square$$

4. Si $v_j = \sum_{k \neq j} a^k v_k$, aleshores $D(v_1, \ldots, v_j, \ldots, v_n) = 0$.

   En efecte,

$$D(v_1, \ldots, v_j, \ldots, v_n) = \sum_{k \neq j} a^k D(v_1, \ldots, v_k, \ldots, v_n).$$

   A cada sumand v_k ocupa els llocs j i k, i, per tant, el sumand s'anul·la. $\square$

5. $D(v_1, \ldots, v_i + \sum_{k \neq i} a^k v_k, \ldots, v_n) = D(v_1, \ldots, v_i, \ldots, v_n)$.

   És conseqüència immediata de 4. $\square$

6. D està determinada pels valors que pren sobre una base $e_1, \ldots, e_n$ de E.

   En efecte, calculem $D(v_1, \ldots, v_n)$. Si $v_i = \sum_{h=1}^{n} a_i^h e_h$,

$$\begin{aligned} D(v_1, \ldots, v_n) &= D(\sum_h a_1^h e_h, \ldots, \sum_h a_n^h e_h) = \\ &= \sum_{h_1, \ldots, h_n = 1}^{n} D(a_1^{h_1} e_{h_1}, \ldots, a_n^{h_n} e_{h_n}) = \\ &= \sum_{h_1, \ldots, h_n = 1}^{n} a_1^{h_1} \ldots a_n^{h_n} D(e_{h_1}, \ldots, e_{h_n}). \end{aligned}$$

   A $D(e_{h_1}, \ldots, e_{h_n})$, els subíndexs $h_1, \ldots, h_n$ poden prendre valors qualssevol de $\{1, \ldots, n\}$, però el sumand s'anul·larà sempre que dos dels subíndexs siguin iguals. Quedaran només, doncs, els sumands en què $h_1, \ldots, h_n$ siguin els mateixos $1, \ldots, n$ permutats. Designem per h la permutació

$$h = \begin{pmatrix} 1 & \ldots & n \\ h_1 & \ldots & h_n \end{pmatrix};$$

   tenim aleshores

$$\begin{aligned} D(v_1, \ldots, v_n) &= \sum_{h \in \mathcal{S}_n} a_1^{h_1} \ldots a_n^{h_n} D(e_{h_1}, \ldots, e_{h_n}) = \quad \text{(per 2)} \\ &= \sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \ldots a_n^{h_n} D(e_1, \ldots, e_n). \end{aligned}$$

Anomenarem *determinant dels vectors* $v_1, \ldots, v_n$ *en la base* $e_1, \ldots, e_n$ l'element de K

$$\det_{(e_i)}(v_1, \ldots, v_n) = \sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \ldots a_n^{h_n}.$$

Aquest element depèn només de les coordenades dels vectors $v_1, \ldots, v_n$ en la base $e_1, \ldots, e_n$, i compleix la condició que, per a tota n-forma lineal alternada D,

$$D(v_1, \ldots, v_n) = \det_{(e_i)}(v_1, \ldots, v_n)\, D(e_1, \ldots, e_n).$$

Això demostra 6.□

7. Sigui $e_1, \ldots, e_n$ una base de E. Donat $k \in K$, existeix una n-forma lineal alternada D, i només una, tal que $D(e_1, \ldots, e_n) = k$.

DEMOSTRACIÓ: La propietat 6 ens diu que, si existeix, D ha d'ésser tal que

$$D(v_1, \ldots, v_n) = \det_{(e_i)}(v_1, \ldots, v_n)k.$$

Només cal comprovar, doncs, que això és sempre una n-forma lineal alternada. Amb les notacions de 6, tenim

$$\begin{aligned} D(v_1, \ldots, av_i, \ldots, v_n) &= \det_{(e_i)}(v_1, \ldots, av_i, \ldots, v_n)\, k = \\ &= \left(\sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \ldots (a a_i^{h_i}) \ldots a_n^{h_n} \right) k = \\ &= a \left(\sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \ldots a_i^{h_i} \ldots a_n^{h_n} \right) k = \\ &= a\, D(v_1, \ldots, v_i, \ldots, v_n), \end{aligned}$$

que demostra la condició (a). Comprovem (b):

$$\begin{aligned} D(v_1, \ldots, v_i + w_i, \ldots, v_n) &= \det_{(e_i)}(v_1, \ldots, v_i + w_i, \ldots, v_n)k = \\ &= \left(\sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \ldots (a_i^{h_i} + b_i^{h_i}) \ldots a_n^{h_n} \right) k = \\ &= \left(\sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \ldots a_i^{h_i} \ldots a_n^{h_n} + \sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \ldots b_i^{h_i} \ldots a_n^{h_n} \right) k = \\ &= D(v_1, \ldots, v_i, \ldots, v_n) + D(v_1, \ldots, w_i, \ldots, v_n). \end{aligned}$$

Resta, finalment, demostrar (c). Suposem que $v_i = v_j$; per tant, llurs coordenades són iguals, $a_i^k = a_j^k$ per a tot k. Aleshores,

$$\det_{(e_i)}(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = \sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \ldots a_i^{h_i} \ldots a_j^{h_j} \ldots a_n^{h_n}.$$

Considerem un sumand qualsevol

$$\varepsilon(\sigma)a_1^{\sigma_1}\dots a_i^{\sigma_i}\dots a_j^{\sigma_j}\dots a_n^{\sigma_n},$$

i comparem-lo amb el sumand corresponent a $\tau = \sigma \circ (i,j)$,

$$\begin{aligned}\varepsilon(\tau)a_1^{\tau_1}\dots a_i^{\tau_i}\dots a_j^{\tau_j}\dots a_n^{\tau_n} &= -\varepsilon(\sigma)a_1^{\sigma_1}\dots a_i^{\sigma_j}\dots a_j^{\sigma_i}\dots a_n^{\sigma_n} = \\ &= -\varepsilon(\sigma)a_1^{\sigma_1}\dots a_j^{\sigma_j}\dots a_i^{\sigma_i}\dots a_n^{\sigma_n}.\end{aligned}$$

Aquests dos sumands sumen 0 i podem suprimir-los en el sumatori. Aquest procés es pot repetir tantes vegades com sigui necessari; al final quedarà

$$\det_{(e_i)}(v_1,\dots,v_i,\dots,v_j,\dots,v_n) = 0,$$

d'on

$$D(v_1,\dots,v_i,\dots,v_j,\dots,v_n) = 0$$

si $v_i = v_j$. □

Així doncs, existeixen tantes n-formes lineals alternades com elements al cos K. Hi ha una manera més maca i precisa d'expressar aquest fet. És la següent: considerem el conjunt $\mathcal{A}(E)$ de les n-formes lineals alternades i, en ell, les operacions donades per

$$\begin{aligned}(D_1 + D_2)(v_1,\dots,v_n) &= D_1(v_1,\dots,v_n) + D_2(v_1,\dots,v_n)\\ (aD)(v_1,\dots,v_n) &= a\,D(v_1,\dots,v_n), \qquad \forall a \in K.\end{aligned}$$

És molt fàcil provar que si D_1, D_2 i D són de $\mathcal{A}(E)$, $D_1 + D_2$ i aD també. Es veu també sense dificultat que $\mathcal{A}(E)$, amb aquestes operacions, és un espai vectorial sobre el cos K.

Sigui ara $e_1,\dots,e_n$ una base de E. L'aplicació

$$\begin{array}{ccc}\mathcal{A}(E) & \longrightarrow & K\\ D & \longmapsto & D(e_1,\dots,e_n)\end{array}$$

és lineal i, per 7, bijectiva. Tenim, doncs, un isomorfisme

$$\mathcal{A}(E) \cong K$$

i, en particular, $\mathcal{A}(E)$ és de dimensió 1. La n-forma corresponent a l'1 de K és precisament

$$\begin{array}{rccc}\det_{(e_i)}: & \overbrace{E\times\cdots\times E}^{n} & \longrightarrow & K\\ & (v_1,\dots,v_n) & \longmapsto & \det_{(e_i)}(v_1,\dots,v_n).\end{array}$$

El valor d'una n-forma D no idènticament zero sobre n vectors linealment independents és, per 6, diferent de zero. El valor de D sobre vectors linealment dependents és sempre zero (per 4). Agafant, en particular, $D = \det_{(e_i)}$, tenim

Proposició 1.1 *$v_1, \ldots, v_n$ són linealment independents si i només si*

$$\det_{(e_i)}(v_1, \ldots, v_n) \neq 0. \square$$

Proposició 1.2 *Siguin $e_1, \ldots, e_n$ i $u_1, \ldots, u_n$ dues bases d'un espai vectorial E. Aleshores,*

$$\det_{(u_i)}(v_1, \ldots, v_n) = \det_{(e_i)}(v_1, \ldots, v_n) \det_{(u_i)}(e_1, \ldots, e_n).$$

DEMOSTRACIÓ: Per a qualsevol n-forma lineal alternada D tenim

$$\begin{aligned} D(v_1, \ldots, v_n) &= \det_{(e_i)}(v_1, \ldots, v_n) D(e_1, \ldots, e_n) = \\ &= \det_{(e_i)}(v_1, \ldots, v_n) \det_{(u_i)}(e_1, \ldots, e_n) D(u_1, \ldots, u_n), \end{aligned}$$

i també $D(v_1, \ldots, v_n) = \det_{(u_i)}(v_1, \ldots, v_n) D(u_1, \ldots, u_n)$.

Podem agafar D tal que $D(u_1, \ldots, u_n) \neq 0$ i obtenim la igualtat desitjada. $\square$

VI.2 Determinant d'una matriu

Donada una matriu $n \times n$ sobre K

$$A = \begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & & \vdots \\ a_1^n & \ldots & a_n^n \end{pmatrix}$$

anomenarem *determinant de A* l'element de K

$$\det A = \sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \cdots a_n^{h_n}.$$

Farem servir també la notació

$$\det A = \begin{vmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & & \vdots \\ a_1^n & \ldots & a_n^n \end{vmatrix}.$$

Fixem un espai vectorial E sobre K i una base $e_1, \ldots, e_n$ de E. Podem interpretar les columnes de A com les coordenades de vectors de E

$$a_i = \sum_{j=1}^{n} a_i^j e_j.$$

Tenim així una correspondència bijectiva entre matrius $n \times n$ i n-ples de vectors de E:

$$\begin{array}{ccc} M_{n \times n}(K) & \longrightarrow & \overbrace{E \times \cdots \times E}^{n} \\ A = (a_i^j) & \longmapsto & (a_1, \ldots, a_n) \end{array}$$

de forma que

$$\det A = \det_{(e_i)}(a_1, \ldots, a_n).$$

Això ens permet traduir les propietats dels determinants de n vectors en propietats dels determinants de les matrius. Per exemple, la condició (a) de n-forma lineal alternada ens diu que si multipliquem els termes d'una columna per un element a de K, el valor del determinant queda multiplicat per a; la propietat 4 diu que si una columna és "combinació lineal" de les altres, el determinant de la matriu és 0. I així totes.

La proposició següent ens dóna una propietat que no s'obté com a "traducció" de cap propietat dels determinants de n vectors.

Proposició 2.1 *Sigui A una matriu $n \times n$ i A^t la seva transposada; aleshores*

$$\det A = \det A^t.$$

DEMOSTRACIÓ: Siguin $A=(a_i^j)$ i $A^t=(b_i^j)$, de forma que $a_i^j=b_j^i$.Tenim, aleshores,

$$\begin{aligned}
\det A &= \sum_{h\in\mathcal{S}_n} \varepsilon(h) a_1^{h(1)} \ldots a_n^{h(n)} = \quad \text{(ordenant els superíndexs)} \\
&= \sum_{h\in\mathcal{S}_n} \varepsilon(h) a^1_{h^{-1}(1)} \ldots a^n_{h^{-1}(n)} = \quad (\text{fent } \sigma = h^{-1}) \\
&= \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma) a^1_{\sigma(1)} \ldots a^n_{\sigma(n)} = \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma) b_1^{\sigma(1)} \ldots b_n^{\sigma(n)} = \\
&= \det A^t. \square
\end{aligned}$$

Aquesta proposició té com a conseqüència que totes les propietats dels determinants de matrius $n \times n$ referents a llurs columnes donen lloc a propietats referents a llurs files. Per exemple, si multipliquem els elements d'una fila per un element a de K, el valor del determinant queda multiplicat per a; si una fila és "combinació lineal" de les altres, el determinant és zero; etc.

VI.3 Determinant d'un endomorfisme

Sigui $f : E \longrightarrow E$ un endomorfisme. Per a tota n-forma lineal alternada D, l'aplicació

$$\begin{array}{rccc}
\hat{f}(D): & \overbrace{E \times \cdots \times E}^{n} & \longrightarrow & K \\
& (v_1, \ldots, v_n) & \longmapsto & D(f(v_1), \ldots, f(v_n))
\end{array}$$

és una n-forma lineal alternada. Tenim, doncs, una aplicació

$$\begin{array}{rccc}
\hat{f}: & \mathcal{A}(E) & \longrightarrow & \mathcal{A}(E) \\
& D & \longmapsto & \hat{f}(D)
\end{array}$$

que resulta ésser lineal. Ara bé, $\mathcal{A}(E)$ és un espai vectorial de dimensió 1 i tota aplicació lineal de $\mathcal{A}(E)$ en ell mateix és una homotècia (V.5.1). En particular, $\hat{f} = a\,I$, on I indica l'aplicació identitat i $a \in K$ és la raó de l'homotècia. Anomenarem *determinant de l'endomorfisme* f la raó de l'homotècia $\hat{f}$

$$\hat{f} = (\det f)\, I.$$

Per calcular explícitament $\det f$ considerem una base $e_1, \ldots, e_n$ de E i una n-forma lineal alternada $D \neq 0$. $\hat{f} = (\det f) I$ implica $\hat{f}(D) = (\det f)\, D$, i, per tant,

$$\hat{f}(D)\,(e_1, \ldots, e_n) = (\det f)\, D(e_1, \ldots, e_n).$$

Per la definició de $\hat{f}$, però,

$$\begin{aligned} \hat{f}(D)(e_1, \ldots, e_n) &= D(f(e_1), \ldots, f(e_n)) = \\ &= \det_{(e_i)}(f(e_1), \ldots, f(e_n)) D(e_1, \ldots, e_n). \end{aligned}$$

Igualant i tenint en compte que $D(e_1, \ldots, e_n) \neq 0$, resulta que

$$\det f = \det_{(e_i)}(f(e_1), \ldots, f(e_n)).$$

La matriu A que té per columnes les coordenades de les imatges $f(e_1), \ldots, f(e_n)$ en la base $e_1, \ldots, e_n$ és la matriu associada a f en la base $e_1, \ldots, e_n$; per tant,

$$\det f = \det A.$$

Un cop arribats en aquest punt, cal preguntar-se per què hem escollit un camí tan "complicat" per definir el determinant d'un endomorfisme i no ens hem limitat a dir que és el determinant de la seva matriu associada. Hi ha diverses raons. La primera és que la definició donada és molt curiosa i elegant; i això és important perquè, sovint, com aquí, un raonament així permet veure la connexió que hi ha entre coses aparentment molt diferents. La segona raó és que tal com ho hem fet ha quedat ben clar que totes les matrius associades a f, en diferents bases, tenen el mateix determinant. Naturalment, aquest fet pot demostrar-se directament, però els càlculs necessaris són molt més "complicats" que la definició donada; la demostració es basa en el fet que el determinant d'un producte de matrius és el producte dels determinants de les matrius (intenteu fer la demostració directa!). La tercera raó és que aquesta definició que hem donat ens permetrà demostrar que $\det AB = \det A \cdot \det B$ sense cap càlcul.

Proposició 3.1 *Si f, g són dos endomorfismes de E,*

$$\widehat{g \circ f} = \hat{f} \circ \hat{g} \qquad i \qquad \hat{I}_E = I_{\mathcal{A}(E)},$$

on I_E i $I_{\mathcal{A}(E)}$ són la identitat a E i a $\mathcal{A}(E)$, respectivament.

DEMOSTRACIÓ: Per a tota $D \in \mathcal{A}(E)$ i $v_1, \ldots, v_n$ de E,

$$\begin{aligned}\widehat{g \circ f}(D)(v_1, \ldots, v_n) &= D(g(f(v_1)), \ldots, g(f(v_n))) = \\ &= \hat{g}(D)(f(v_1), \ldots, f(v_n)) = \\ &= \hat{f}(\hat{g}(D))(v_1, \ldots, v_n) = (\hat{f} \circ \hat{g})(D)(v_1, \ldots, v_n),\end{aligned}$$

d'on $\widehat{g \circ f} = \hat{f} \circ \hat{g}$.
$\hat{I}_E(D)(v_1, \ldots, v_n) = D(I_E(v_1), \ldots, I_E(v_n)) =$
$= D(v_1, \ldots, v_n) = I_{\mathcal{A}(E)}(D)(v_1, \ldots, v_n)$,
d'on $\hat{I}_E = I_{\mathcal{A}(E)}$.□

Corol·lari 3.2 *Si f, g són dos endomorfismes de E i I és la identitat, es compleix*

$$\det(g \circ f) = \det f \cdot \det g \qquad i \qquad \det I = 1.$$

Si f és bijectiva,

$$\det f^{-1} = (\det f)^{-1}.\square$$

Corol·lari 3.3 *Si A, B són de $M_{n\times n}(K)$ i I és la matriu identitat, aleshores*

$$\det AB = \det A \cdot \det B \qquad i \qquad \det I = 1.\square$$

De (3.2) es dedueix que els automorfismes tenen determinant no nul; el recíproc també es cert.

Proposició 3.4 *Un endomorfisme f de E és automorfisme si i només si* $\det f \neq 0$.

DEMOSTRACIÓ: Sigui $e_1, \ldots, e_n$ una base de E. f és automorfisme si i només si els vectors $f(e_1), \ldots, f(e_n)$ són linealment independents, que, per (1.1), equival a

$$\det_{(e_i)}(f(e_1), \ldots, f(e_n)) \neq 0;$$

és a dir, $\det f \neq 0$.□

Recordem, finalment, que, si A és la matriu associada a un endomorfisme f en una certa base, la transposada A^t és la matriu associada a l'aplicació dual f' en la base dual de l'anterior (V.6). Això, juntament amb (2.1), ens diu que f i f' tenen el mateix determinant.

Proposició 3.5 *Sigui f un endomorfisme i f' el seu dual. Aleshores* $\det f = \det f'$.□

VI.4 Regla de Laplace

En aquest apartat anem a donar una altra expressió del determinant d'una matriu que permet, moltes vegades, calcular més còmodament aquest determinant. També en deduirem una manera de calcular la matriu inversa de A. Necessitem, però, una notació adient.

$C_p(1,2,\ldots,n)$, o simplement C_p, indicarà el conjunt de tots els subconjunts de p elements de $\{1,2,\ldots,n\}$. Si $H \in C_p$, H' indicarà el complementari de H, és a dir, el conjunt d'elements de $\{1,2,\ldots,n\}$ que no són a H. f_H designarà la permutació que compleix

$$\{f_H(1),\ldots,f_H(p)\} \underset{f_H(1)<\ldots<f_H(p)}{} = H \qquad \{f_H(p+1),\ldots,f_H(n)\} \underset{f_H(p+1)<\ldots<f_H(n)}{} = H'.$$

Un *menor d'ordre* p d'una matriu $A = (a_i^j)$ és una matriu $p \times p$ formada pels elements de A situats en p files i p columnes prefixades. És a dir, per a cada elecció de p files $H = \{i_1,\ldots,i_p\}$ i p columnes $L = \{j_1,\ldots,j_p\}$, hi ha un menor d'ordre p, que és

$$\begin{pmatrix} a_{j_1}^{i_1} & \ldots & a_{j_p}^{i_1} \\ \vdots & & \vdots \\ a_{j_1}^{i_p} & \ldots & a_{j_p}^{i_p} \end{pmatrix}.$$

Denotarem per A_L^H el determinant d'aquest menor.

Proposició 4.1 (Regla de Laplace) *Fixem $L \in C_p(1,2,\ldots,n)$. Llavors,*

$$\det A = \sum_{H \in C_p} \varepsilon(f_L)\,\varepsilon(f_H)\,A_L^H A_{L'}^{H'}.$$

DEMOSTRACIÓ: Tenim

$$\begin{aligned} \det A &= \sum_{h \in \mathcal{S}_n} \varepsilon(h) a_1^{h_1} \ldots a_n^{h_n} = \text{(reordenant el producte en cada sumand)} \\ &= \sum_{h \in \mathcal{S}_n} \varepsilon(h) a_{f_L(1)}^{h f_L(1)} \ldots a_{f_L(n)}^{h f_L(n)} = \quad \text{(posant } g = h f_L\text{)} \\ &= \sum_{g \in \mathcal{S}_n} \varepsilon(g)\,\varepsilon(f_L)\, a_{f_L(1)}^{g(1)} \ldots a_{f_L(n)}^{g(n)}. \end{aligned}$$

Per a cada $g \in \mathcal{S}_n$ sigui $H = \{g(1),\ldots,g(p)\}$. Aleshores $g = \sigma' \circ \sigma \circ f_{H'}$ on σ permuta els elements de H i deixa fixos els de H' i σ' permuta els elements de H' i deixa fixos els de H. L'última expressió de $\det A$ trobada es pot escriure així:

$$\sum_{\substack{H \in C_p \\ \sigma,\sigma'}} \varepsilon(f_L)\,\varepsilon(\sigma')\,\varepsilon(\sigma)\,\varepsilon(f_H)\, a_{f_L(1)}^{\sigma'\sigma f_H(1)} \ldots a_{f_L(n)}^{\sigma'\sigma f_H(n)}.$$

Observem que si $j \in \{1,\dots,p\}$, $\sigma'\sigma f_H(j) = \sigma f_H(j)$, i, si $j \in \{p+1,\dots,n\}$, $\sigma'\sigma f_H(j) = \sigma' f_H(j)$. Podem reordenar, doncs, l'expressió anterior, obtenint

$$\sum_{H\in C_p} \varepsilon(f_L)\,\varepsilon(f_H) \left(\sum_{\sigma} \varepsilon(\sigma) a_{f_L(1)}^{\sigma f_H(1)} \dots a_{f_L(p)}^{\sigma f_H(p)}\right) \cdot \left(\sum_{\sigma'} \varepsilon(\sigma') a_{f_L(p+1)}^{\sigma' f_H(p+1)} \dots a_{f_L(n)}^{\sigma' f_H(n)}\right).$$

En el sumatori del primer parèntesi apareixen només les columnes $\{f_L(1),\dots, f_L(p)\} = L$ i les files $\{f_H(1),\dots,f_H(p)\} = H$ permutades de totes les maneres possibles. Aquest sumatori és, per tant, l'expressió del determinant del menor format per les columnes L i les files H: A_L^H. De la mateixa manera, el sumatori del segon parèntesi resulta ésser $A_{L'}^{H'}$. Hem obtingut, doncs,

$$\det A = \sum_{H\in C_p} \varepsilon(f_L)\,\varepsilon(f_H)\, A_L^H\, A_{L'}^{H'}. \square$$

Corol·lari 4.2 $\det A = \sum_{i=1}^{n} (-1)^{i+j} a_j^i A_{j'}^{i'}$, *on* $i' = \{i\}'$, $j' = \{j\}'$.

DEMOSTRACIÓ: Apliquem la regla de Laplace (4.1) per a $L = \{j\}$. Per a cada $H = \{i\}$, $A_L^H = a_j^i$ i $A_{L'}^{H'} = A_{j'}^{i'}$. A més, es veu fàcilment que $\varepsilon(f_L) = (-1)^{j-1}$ i $\varepsilon(f_H) = (-1)^{i-1}$. Substituint ara a (4.1) obtenim l'expressió de l'enunciat. □

L'expressió del corol·lari 4.2 es coneix com el *desenvolupament del determinant pels elements de la columna* j.

Recordem que si A^t és la matriu transposada de A, $\det A^t = \det A$ (2.1). D'aquí que la regla de Laplace sigui vàlida també fixant p files ($H \in C_p$),

$$\det A = \sum_{L\in C_p} \varepsilon(f_H)\,\varepsilon(f_L)\, A_L^H\, A_{L'}^{H'}.$$

Anàlogament, tenim una expressió del *desenvolupament d'un determinant pels elements d'una fila* i

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_j^i A_{j'}^{i'}.$$

Es diu *adjunt d'un element* a_j^i d'una matriu A

$$X_j^i = (-1)^{i+j} A_{j'}^{i'}.$$

Corol·lari 4.3 $\sum_{i=1}^{n} a_j^i X_j^i = \det A.$

Per a $j \neq k$, $\quad \sum_{i=1}^{n} a_j^i X_k^i = 0.$

DEMOSTRACIÓ: La primera afirmació és (4.2). Per demostrar la segona, considerem una matriu $\bar{A}$ amb les mateixes columnes que A, llevat de la columna k on torna a aparèixer la columna j. Clarament, $\det \bar{A} = 0$. Ara bé, si desenvolupem $\bar{A}$ pels termes de la columna k obtenim

$$0 = \det \bar{A} = \sum_{i=1}^{n} \bar{a}_k^i \bar{X}_k^i \ ;$$

però, tal com hem definit $\bar{A}$, $\bar{a}_k^i = a_j^i$ i $\bar{X}_k^i = X_k^i$, d'on s'obté el resultat. □

(4.3) es pot interpretar d'una altra manera. Considerem la matriu $B = (b_i^j)$ amb $b_i^j = X_j^i$; aleshores

$$BA = (\det A)\, I \ ,$$

on I indica la matriu identitat. Si $\det A = 0$, $BA = 0$; si $\det A \neq 0$, aleshores $(\det A)^{-1} BA = I$.

Corol·lari 4.4 *La matriu $C = (c_i^j)$ amb $c_i^j = X_j^i (\det A)^{-1}$ és la inversa de la matriu A.* □

Acabarem aquest apartat aplicant la regla de Laplace a un cas molt típic. Suposem que en una matriu A els elements de les p primeres columnes són 0 llevat, potser, dels elements que ocupen les p primeres files:

$$A = \begin{pmatrix} a_1^1 & \dots & a_p^1 & a_{p+1}^1 & \dots & a_n^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_1^p & \dots & a_p^p & a_{p+1}^p & \dots & a_n^p \\ 0 & \dots & 0 & a_{p+1}^{p+1} & \dots & a_n^{p+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{p+1}^n & \dots & a_n^n \end{pmatrix} .$$

Agafant a (4.1) $E = \{1, \dots, p\}$, s'obté

$$\det A = \begin{vmatrix} a_1^1 & \dots & a_p^1 \\ \vdots & & \vdots \\ a_1^p & \dots & a_p^p \end{vmatrix} \cdot \begin{vmatrix} a_{p+1}^{p+1} & \dots & a_n^{p+1} \\ \vdots & & \vdots \\ a_{p+1}^n & \dots & a_n^n \end{vmatrix} .$$

Un cas particular, per a $p = 1$, és el següent:

$$\begin{vmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ 0 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ 0 & a_2^n & \dots & a_n^n \end{vmatrix} = a_1^1 \begin{vmatrix} a_2^2 & \dots & a_n^2 \\ \vdots & & \vdots \\ a_2^n & \dots & a_n^n \end{vmatrix} .$$

Així, per exemple, si tots els elements sota la "diagonal principal" són 0 ($a_i^j = 0$ quan $i < j$), tenim

$$\begin{vmatrix} a_1^1 & a_2^1 & a_3^1 & \dots & a_n^1 \\ 0 & a_2^2 & a_3^2 & \dots & a_n^2 \\ 0 & 0 & a_3^3 & \dots & a_n^3 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_n^n \end{vmatrix} = a_1^1 a_2^2 \dots a_n^n.$$

El mètode general més pràctic per a calcular un determinant és reduir-lo a un determinant "triangular" d'aquest tipus, aplicant 2 i 5 de (VI.1) reiteradament. El procediment és anàleg al que explicarem en detall a (VII.5).

VI.5 Càlcul del rang d'una matriu

Sigui E un espai vectorial de dimensió n. A (1.1) hem donat un criteri per a saber si n vectors de E són, o no, linealment independents: ho són si i només si el seu determinant és diferent de 0. En aquest apartat volem ampliar aquest criteri per poder reconèixer si k vectors $v_1, \dots, v_k$ de E són, o no, linealment independents, quins són combinació lineal de la resta i quines són les combinacions lineals que els relacionen.

Sigui $e_1, \dots, e_n$ una base de E. Un cop fixada una base, cada vector vindrà representat per una n-pla d'elements del cos K, les seves coordenades, i cada conjunt de vectors $v_1, \dots, v_k$ per una matriu de k columnes formades per les coordenades dels k vectors. Tota matriu A, $n \times k$, correspon a k vectors de E; anomenarem *rang de A* el nombre de vectors-columna de A linealment independents. El problema plantejat equival, doncs, al càlcul del rang d'una matriu.

Proposició 5.1 *El rang d'una matriu A és el més gran dels ordres dels menors de A amb determinant no nul.*

DEMOSTRACIÓ: Indicarem per a_j el vector corresponent a la columna j de $A = (a_j^i)$; és a dir, les coordenades de a_j en la base prefixada són $(a_j^1, \dots, a_j^n)$. Posem $r = \operatorname{rang} A$.

Demostrarem, primer, que tot menor d'ordre $p > r$ té determinant zero. Siguin $j_1, \dots, j_p$ les columnes amb les quals s'ha format el menor. Els vectors $a_{j_1}, \dots, a_{j_p}$ són linealment dependents i, per tant, un d'ells és combinació lineal de la resta. Amb més raó, en el menor considerat una de les columnes serà combinació lineal de la resta i el determinant serà zero.

Vegem ara que hi ha un menor d'ordre r amb determinant no nul. A la matriu A hi ha r columnes linealment independents; siguin $a_{j_1}, \dots, a_{j_r}$. Completem aquests vectors amb vectors de la base $e_1, \dots, e_n$ en què treballem fins a obtenir una base de

l'espai E: $a_{j_1},\ldots,a_{j_r},e_{h_{r+1}},\ldots,e_{h_n}$. El determinant d'aquests n vectors és diferent de zero:

$$0 \neq \det_{(e_i)}(a_{j_1},\ldots,a_{j_r},e_{h_{r+1}},\ldots,e_{h_n}) = \begin{vmatrix} a^1_{j_1} & \ldots & a^1_{j_r} & 0 & 0 & \ldots & 0 \\ . & & . & . & . & & . \\ . & & . & . & . & & . \\ . & & . & 1 & . & & . \\ . & & . & . & . & & 1 \\ . & & . & . & 1 & & . \\ . & & . & . & . & & . \\ a^n_{j_1} & \ldots & a^n_{j_r} & 0 & 0 & \ldots & 0 \end{vmatrix}.$$

A cada una de les $n-r$ últimes columnes tots els elements són 0 llevat d'un que val 1; aquest element apareix a cada columna en una fila diferent.

Apliquem, ara, la regla de Laplace al càlcul d'aquest determinant. Fixem, per a això, les $n-r$ últimes columnes; l'únic menor de determinant no nul que hi podem formar és el corresponent a les $n-r$ files en què apareix un 1. Si $i_1,\ldots,i_r$ són les files restants, queda

$$0 \neq \det_{(e_i)}(a_{j_1},\ldots,e_{h_n}) = \pm \begin{vmatrix} a^{i_1}_{j_1} & \ldots & a^{i_1}_{j_r} \\ \vdots & & \vdots \\ a^{i_r}_{j_1} & \ldots & a^{i_r}_{j_r} \end{vmatrix}.$$

Obtenim així un menor d'ordre r amb determinant no nul.□

Corol·lari 5.2 *Una matriu A i la seva transposada A^t tenen el mateix rang:*

$$\operatorname{rang} A = \operatorname{rang} A^t.\square$$

Corol·lari 5.3 *Una aplicació lineal f i la seva dual f' tenen el mateix rang.*

DEMOSTRACIÓ: El rang d'una aplicació lineal $f : E \longrightarrow F$ és la dimensió de $\operatorname{Im} f$ (V.1). Si $e_1,\ldots,e_n$ és una base de E, aleshores $\operatorname{Im} f = \langle f(e_1),\ldots,f(e_n)\rangle$. Les coordenades de $f(e_1),\ldots,f(e_n)$ formen les columnes de la matriu A associada a f. Per tant,

$$\operatorname{rang} f = \operatorname{rang} A.$$

El resultat es dedueix ara de (5.2) i de (V.6.2).□

Nota:

El corol·lari 5.3 s'obté també com a conseqüència de (V.7.1).

La proposició 5.1 dóna un mètode per a calcular el rang d'una matriu A mitjançant el càlcul dels determinants dels menors de A, començant pels d'ordre màxim fins a trobar-ne un de no nul. La proposició següent redueix sensiblement el nombre de determinants a calcular, en moltes ocasions.

Proposició 5.4 *Siguin* $a_i = (a_i^1, \ldots, a_i^n)$, $i = 1, \ldots, r$, *vectors linealment independents, i sigui*

$$M = \begin{vmatrix} a_1^{i_1} & \ldots & a_r^{i_1} \\ \vdots & & \vdots \\ a_1^{i_r} & \ldots & a_r^{i_r} \end{vmatrix} \neq 0.$$

Un vector $v = (v^1, \ldots, v^n)$ *és combinació lineal de* $a_1, \ldots, a_r$ *si i només si, per a tot* j,

$$\begin{vmatrix} a_1^{i_1} & \ldots & a_r^{i_1} & v^{i_1} \\ \vdots & & \vdots & \vdots \\ a_1^{i_r} & \ldots & a_r^{i_r} & v^{i_r} \\ a_1^{j} & \ldots & a_r^{j} & v^{j} \end{vmatrix} = 0.$$

DEMOSTRACIÓ: El fet que, si $a_1, \ldots, a_r, v$ són linealment dependents, aquests determinants siguin 0, és conseqüència de (5.1). Suposem, recíprocament, que aquests determinants són tots zero. Desenvolupant per l'última fila, obtenim

$$a_1^j \begin{vmatrix} a_2^{i_1} & \ldots & a_r^{i_1} & v^{i_1} \\ \vdots & & \vdots & \vdots \\ a_2^{i_r} & \ldots & a_r^{i_r} & v^{i_r} \end{vmatrix} - a_2^j \begin{vmatrix} a_1^{i_1} & a_3^{i_1} & \ldots & a_r^{i_1} & v^{i_1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^{i_r} & a_3^{i_r} & \ldots & a_r^{i_r} & v^{i_r} \end{vmatrix} + \cdots$$

$$\cdots \pm v^j \begin{vmatrix} a_1^{i_1} & \ldots & a_r^{i_1} \\ \vdots & & \vdots \\ a_1^{i_r} & \ldots & a_r^{i_r} \end{vmatrix} = 0 \qquad \text{per a tot } j.$$

Denotem per $M_1, M_2, \ldots$ els determinants d'aquesta expressió. L'últim és M i cap d'ells no depèn de l'índex j. Aïllant v^j obtenim

$$v^j = \pm M_1 M^{-1} a_1^j \pm M_2 M^{-1} a_2^j \pm \cdots \pm M_r M^{-1} a_r^j \qquad \text{per a tot } j,$$

d'on

$$v = \pm (M_1 M^{-1}) a_1 \pm (M_2 M^{-1}) a_2 \pm \cdots \pm (M_r M^{-1}) a_r \,,$$

que ens dóna v com a combinació lineal de $a_1, \ldots, a_r$. □

Nota:

Observem que a (5.4) la condició és òbvia per a $j = i_1, \ldots, i_r$.

Exemple:

Considerem la matriu

$$A = \begin{pmatrix} 1 & -1 & 1 & 1 \\ -1 & -2 & 1 & 1 \\ 4 & 3 & 6 & -1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

$a_1 = (1,-1,4,0)$ i $a_2 = (1,-2,3,0)$ són clarament linealment independents, ja que llurs coordenades no són proporcionals. Escollim

$$M = \begin{vmatrix} 1 & 1 \\ -1 & -2 \end{vmatrix} = -1 \neq 0.$$

Per veure si $a_3 = (1,1,6,0)$ depèn o no linealment de a_1, a_2, hem de calcular

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 4 & 3 & 6 \end{vmatrix} = 0 \qquad \begin{vmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

Així doncs, a_3 és combinació lineal de a_1, a_2. Per trobar aquesta combinació fem

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ a_1^j & a_2^j & a_3^j \end{vmatrix} = 3a_1^j - 2a_2^j - a_3^j = 0,$$

d'on

$$a_3^j = 3a_1^j - 2a_2^j \qquad \text{per a tot } j,$$

i, per tant,

$$a_3 = 3a_1 - 2a_2.$$

Continuem calculant el rang de A estudiant si $a_4 = (1,1,-1,2)$ és o no combinació lineal de a_1, a_2. Per a això, calculem els determinants dels menors d'ordre 3 formats a partir de M. Tenim

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 0 & 0 & 2 \end{vmatrix} = -2 \neq 0.$$

a_1, a_2, a_4 són, doncs, linealment independents i la matriu A té rang igual a 3.

VI.6 Nota històrica

Els determinants apareixen per primer cop en la resolució de sistemes d'equacions lineals (veure Cap. VII) el 1772 de la mà d'Alexandre-Théophile Vandermonde

(1735–1796) i s'apliquen ja el segle 19 a la teoria de l'eliminació, transformació de coordenades, canvi de variable, etc.

La paraula "determinant" fou introduïda per Carl Friedrich Gauss (1777–1855) en l'estudi de certes formes quadràtiques, però el tractament sistemàtic i gairebé actual és degut a Augustin-Louis Cauchy (1789–1857) l'any 1815, que prova entre altres propietats la regla de Laplace (demostrada ja per Pierre-Simon de Laplace (1749–1885) el 1825), demostrant pràcticament totes les propietats esmentades en el present capítol, i a James Joseph Sylvester (1817–1897), que l'aplica a problemes de la teoria d'equacions.

Leopold Kronecker (1823–1891) i Karl Wilhelm Weierstrass (1815–1897) (segons alguns el millor professor que mai no hagi tingut una Universitat) introduïren en llurs cursos a Berlín els determinants com a formes multilineals alternades.

VI.7 Exercicis

1. Calculeu el determinant de la matriu $A = (a^i_j)$, on $a^i_j = |i - j|$.

2. Proveu que $(x-1)^3$ divideix el polinomi

$$\begin{vmatrix} 1 & x & x^2 & x^3 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \end{vmatrix}.$$

3. Donada la matriu

$$M = \left(\begin{array}{c|c} 0 & A \\ \hline \operatorname{adj} A & 0 \end{array}\right),$$

calculeu $\det M$ en funció de $\det A$. ($\operatorname{adj} A$ indica la matriu adjunta de la A, que s'obté a partir de A substituint cada element pel seu adjunt).

4. Una matriu $n \times n$ $A = (a^j_i)$ es diu *hemisimètrica* si $a^j_i = -a^i_j$ per a tot i, j. Proveu que si A és hemisimètrica $\det A = (-1)^n \det A$. Deduïu-ne que les matrius hemisimètriques d'ordre senar tenen determinant zero.

5. Descomponeu el polinomi

$$\begin{vmatrix} 1+x+x^2 & 1 & 1 & 1 & 1 \\ 1 & 1+x+x^2 & 1 & 1 & 1 \\ 1 & 1 & 1+x+x^2 & 1 & 1 \\ 1 & 1 & 1 & 1+x+x^2 & 1 \\ 1 & 1 & 1 & 1 & 1+x+x^2 \end{vmatrix} \in \mathbf{C}[x]$$

en factors irreductibles.

6. Sigui E l'espai vectorial de les funcions reals de variable real generat per sin i cos. Calculeu el determinant de l'endomorfisme de E definit per la derivació.

7. Demostreu que l'aplicació que va de les matrius invertibles de $M_{n\times n}(K)$(que denotarem $GL(n,K)$) al grup multiplicatiu del cos K:

$$\det : GL(n,K) \longrightarrow K - \{0\}$$

és un homomorfisme de grups. Estudieu-lo.

8. Sigui A una matriu $n \times n$ i $\operatorname{adj} A$ la matriu adjunta de A. Demostreu:

a) $\det(\operatorname{adj} A) = (\det A)^{n-1}$.

b) Si $\operatorname{rang} A = n - 1$, $\operatorname{rang}(\operatorname{adj} A) = 1$.

9. Torneu a fer ara els exercicis 13 i 14 del capøitol IV.

10. Considerem el determinant

$$\mathrm{D}(n,k) = \begin{vmatrix} 1^k & 2^k & \dots & n^k \\ 2^k & 3^k & \dots & (n+1)^k \\ \vdots & \vdots & & \vdots \\ n^k & (n+1)^k & \dots & (2n-1)^k \end{vmatrix}.$$

a) Calculeu $\mathrm{D}(1,1)$, $\mathrm{D}(2,1)$, $\mathrm{D}(3,1)$, $\mathrm{D}(4,1)$.

b) Demostreu que $\mathrm{D}(n,2) = 0 \quad \forall n > 3$.

c) Demostreu que $\mathrm{D}(n,k) = 0 \quad \forall n > k+1$.

11. Demostreu que si A és invertible $(A^{-1})^{\mathrm{t}} = (A^{\mathrm{t}})^{-1}$.

VI.8 Exercicis de programar

12. Donada una matriu $A \in M_{n\times n}(\mathbf{R})$, elaboreu un programa que calculi $\det A$ pel mètode explicat al final del §4.

(Indicació: permutant files, si cal, aconseguireu $a_1^1 \neq 0$. Guardeu en una variable el possible canvi de signe. Anul·leu tots els elements de la primera columna sota la diagonal. Si X_1^1 és l'adjunt de a_1^1, aleshores $\det A = a_1^1 X_1^1$. Repetiu el procés per a X_1^1).

13. Compteu quantes sumes i multiplicacions cal fer per obtenir $\det A$ pel mètode de l'exercici anterior. Compteu també quantes en caldria fer utilitzant el desenvolupament de 4.2. Observeu la gran diferència que hi ha quan n és gran i traieu-ne conseqüències.

14. (Veure l'exercici IV.17). Demostreu que si $\det A \neq 0$ la descomposició LU és sempre possible, llevat potser d'haver de permutar les files de A. Observeu que la descomposició $A = LU$, si és possible, és única.

15. Apliqueu l'exercici anterior per

a) calcular $\det A$.

b) resoldre un sistema d'equacions $Ax = b$.

(Compteu quantes sumes i multiplicacions calen per a resoldre $Ax = b$ per aquest mètode. Observeu que és essencialment equivalent al mètode de Gauss (Cap.VII), sempre que $\det A \neq 0$).

Capítol VII

Sistemes d'equacions lineals

VII.1 Planteig del problema

Volem resoldre el següent problema: suposem que tenim un sistema d'equacions lineals

$$\left\{\begin{array}{rcl} a_1^1 x^1 + \ldots + a_n^1 x^n & = & b^1 \\ \ldots\ldots\ldots\ldots & & \\ a_1^m x^1 + \ldots + a_n^m x^n & = & b^m \end{array}\right.$$

on els a_i^j i b^j són elements coneguts d'un cos K. Es tracta de trobar les n-ples $(x^1, \ldots, x^n)$ de K^n que satisfan totes aquestes equacions. Posem

$$A = \begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \ldots & a_n^m \end{pmatrix} \quad b = \begin{pmatrix} b^1 \\ \vdots \\ b^m \end{pmatrix} \quad x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

Podem escriure llavors el sistema anterior així:

$$Ax = b.$$

Aquest mateix problema es pot plantejar d'una altra forma. Siguin E i F dos espais vectorials sobre K amb bases $u_1, \ldots, u_n$ i $v_1, \ldots, v_m$ respectivament. Sabem que existeix una aplicació lineal $f : E \longrightarrow F$ amb matriu associada en aquestes bases la matriu A (V.2). Designem per b el vector $b = b^1 v_1 + \ldots + b^m v_m$ de F i per x el vector $x = x^1 u_1 + \ldots + x^n u_n$ de E. La condició $Ax = b$ equival a

$$f(x) = b.$$

El problema consisteix, per tant, a trobar les antiimatges del vector b per l'aplicació f.

Recíprocament, si suposem donada una aplicació lineal $f : E \longrightarrow F$ i un vector $b \in F$, el problema de trobar les antiimatges x de b equival a resoldre el sistema d'equacions lineals

$$Ax = b$$

on A és la matriu associada a f en unes certes bases i b i x són matrius d'una columna formades per les coordenades dels vectors corresponents. Tant en la primera interpretació com en l'altra el nostre objectiu és:

1. Saber quan el problema té solució i quan no.
2. Saber quantes solucions té.
3. Donar un mètode per a trobar totes les solucions.

Per resoldre cada una d'aquestes qüestions farem servir la interpretació que ens resulti més còmoda. En general, els raonaments de tipus teòric són més simples en el llenguatge d'aplicacions lineals i la resolució dels casos concrets es fa amb el llenguatge de matrius.

VII.2 Existència de solucions

En el problema plantejat a l'apartat anterior tenim que $\exists x \in E$ amb $f(x) = b$

$$\begin{array}{lll} \Leftrightarrow & b \in \operatorname{Im} f & \Leftrightarrow \quad \langle f(u_1), \ldots, f(u_n) \rangle = \langle f(u_1), \ldots, f(u_n), b \rangle \\ \Leftrightarrow & \operatorname{rang} f & = \quad \dim \operatorname{Im} f = \dim \langle f(u_1), \ldots, f(u_n), b \rangle. \end{array}$$

En la demostració de (VI.5.3) vam veure que $\operatorname{rang} f = \operatorname{rang} A$. Anàlogament,

$$\dim \langle f(u_1), \ldots, f(u_n), b \rangle = \operatorname{rang}(A, b)$$

on (A, b) indica la matriu que s'obté afegint a A una columna formada per les coordenades de b. Així doncs, en llenguatge de matrius tenim que el sistema $Ax = b$ té solució si i només si $\operatorname{rang} A = \operatorname{rang}(A, b)$. A (VI.5) vam donar un mètode per a calcular el rang d'una matriu. Tenim, per tant, resolt el problema de saber si el sistema té o no solucions. Quantes solucions hi ha? És a dir, quantes antiimatges té b? En la demostració del teorema d'isomorfisme (V.3.1) vam veure que tots els vectors de E que s'apliquen en el mateix vector de F formen una classe mòdul el nucli de f : $x_0 + \operatorname{Nuc} f$. Així doncs, b té tantes antiimatges com vectors té $\operatorname{Nuc} f$. A més, totes les antiimatges s'obtenen sumant a una solució particular x_0 els vectors de $\operatorname{Nuc} f$. Els vectors de $\operatorname{Nuc} f$ tenen per coordenades les solucions del sistema homogeni associat al sistema donat

$$Ax = 0.$$

Per tant, les solucions del sistema $Ax = b$ s'obtenen sumant a una solució particular les solucions del sistema homogeni associat.

VII.3 Regla de Cramer

Anem a donar en aquest apartat un mètode per a trobar les solucions d'un sistema d'equacions lineals en un cas molt particular. Suposem que a

$$Ax = b$$

la matriu A és quadrada, és a dir, $n \times n$, i $\det A \neq 0$. O bé, equivalentment, suposem que a $f(x) = b$ f és un isomorfisme. Hi ha, aleshores, una solució i només una, que és

$$x = A^{-1}b.$$

Vam veure a (VI.4.4) que si $A_j^{-1} = (c_j^i)$, $c_j^i = X_i^j(\det A)^{-1}$, on X_i^j és l'adjunt de l'element a_i^j de A. Per tant,

$$x^i = \sum_{j=1}^{n} X_i^j b^j (\det A)^{-1}.$$

Ara bé, $\sum_{j=1}^{n} X_i^j b^j$ és el desenvolupament pels termes de la columna i del determinant d'una matriu amb els mateixos elements que A, llevat de la columna i que ha estat substituïda per $(b^1, \ldots, b^n)$. Designem aquest determinant per $\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)$. Llavors

$$x^i = \frac{\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)}{\det A} \qquad i = 1, \ldots, n.$$

Aquesta fórmula es coneix amb el nom de *regla de Cramer* .

VII.4 Resolució d'un sistema d'equacions lineals

Donat el sistema $Ax = b$,

$$\begin{cases} a_1^1 x^1 + \ldots + a_n^1 x^n &= b^1 \\ \cdots\cdots\cdots\cdots & \\ a_1^m x^1 + \ldots + a_n^m x^n &= b^m, \end{cases}$$

suposem que existeixen solucions, és a dir, que

$$\operatorname{rang} A = \operatorname{rang}(A, b) = r.$$

Reordenant les equacions i les incògnites convenientment podem suposar (VI.5.1) que

$$M = \begin{vmatrix} a_1^1 & \ldots & a_r^1 \\ \vdots & & \vdots \\ a_1^r & \ldots & a_r^r \end{vmatrix} \neq 0.$$

A la matriu (A, b), les $m-r$ últimes files són combinacions lineals de les anteriors; és a dir, en el sistema donat, les $m-r$ últimes equacions són combinacions lineals de les r primeres. Per tant, el sistema original té exactament les mateixes solucions que el sistema

$$\begin{cases} a_1^1 x^1 + \ldots + a_n^1 x^n &= b^1 \\ \cdots\cdots\cdots\cdots & \\ a_1^r x^1 + \ldots + a_n^r x^r &= b^r \end{cases}$$

format per les r primeres equacions. N'hi ha prou, doncs, amb trobar les solucions d'aquest sistema parcial.

Escriurem el sistema anterior en la forma

$$\begin{cases} a_1^1 x^1 + \ldots + a_r^1 x^r &= b^1 - a_{r+1}^1 x^{r+1} - \ldots - a_n^1 x^n \\ \cdots\cdots\cdots\cdots & \\ a_1^r x^1 + \ldots + a_r^r x^r &= b^r - a_{r+1}^r x^{r+1} - \ldots - a_n^r x^n. \end{cases}$$

Per a cada un dels conjunts de valors que donem a $x^{r+1}, \ldots, x^n$ arbitràriament, això és un sistema de r equacions amb r incògnites. El determinant de la seva matriu és $M \neq 0$. Podem aplicar, doncs, la regla de Cramer i obtenim uns valors únics per a les incògnites $x^1, \ldots, x^r$:

$$x^i = M^{-1} \det(a_1, \ldots, a_{i-1}, b - a_{r+1}x^{r+1} - \ldots - a_n x^n, a_{i+1}, \ldots, a_r) \quad i = 1, \ldots, r$$

(Aquí $b = (b^1, \ldots, b^r)$ i $a_j = (a_j^1, \ldots, a_j^r)$).

Ara bé,

$$\det(a_1, \ldots, b - a_{r+1}x^{r+1} - \ldots - a_n x^n, \ldots, a_r) = \det(a_1, \ldots, b, \ldots, a_r) - \\ - \det(a_1, \ldots, a_{r+1}, \ldots, a_r)x^{r+1} - \ldots - \det(a_1, \ldots, a_n, \ldots, a_r)x^n.$$

Observem que aquests determinants s'obtenen substituint a M la columna i per b, a_{r+1}, ..., a_n successivament. Posem

$$M_b^i = M^{-1} \det(a_1, \ldots, b, \ldots a_r), \quad M_j^i = M^{-1} \det(a_1, \ldots, a_j, \ldots, a_r);$$

aleshores

$$x^i = M_b^i - M_{r+1}^i x^{r+1} - \ldots - M_n^i x^n \qquad i = 1, \ldots, r.$$

Aquesta expressió ens dóna la *solució general* del sistema en funció de les $n-r$ incògnites arbitràries $x^{r+1}, \ldots, x^n$.

Observacions:

1. Agafant $x^{r+1} = \ldots = x^n = 0$ obtenim una solució particular amb $x^i = M_b^i$ per a $i = 1, \ldots, r$:

$$(M_b^1, \ldots, M_b^r, 0, \ldots, 0).$$

2. Per resoldre el sistema homogeni $Ax = 0$ associat al nostre, hem de substituir b per $(0,\ldots,0)$ en tot l'anterior. Resulta aleshores

$$x^i = -M^i_{r+1}x^{r+1} - \ldots - M^i_n x^n \quad i = 1,\ldots,r.$$

El conjunt d'aquestes solucions forma un espai vectorial (que correspon a $\mathrm{Nuc} f$ tal com vam veure a l'apartat 1). Podem obtenir una base d'aquest espai de les solucions del sistema homogeni, donant a les incògnites arbitràries $x^{r+1},\ldots,x^n$ valors 0 llevat d'una amb valor 1:

$$(-M^1_{r+1},\ldots,-M^r_{r+1},1,0,\ldots,0)$$

$$(-M^1_{r+2},\ldots,-M^r_{r+2},0,1,\ldots,0)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$(-M^1_n,\ldots,-M^r_n,0,0,\ldots,1).$$

Aquestes solucions són, en efecte, linealment independents i el seu nombre és

$$n - r = n - \mathrm{rang}\, A = n - \dim \mathrm{Im}\, f = \dim \mathrm{Nuc}\, f.$$

3. La solució general obtinguda és suma de la solució particular

$$(M^1_b,\ldots,M^r_b,0,\ldots,0)$$

i la solució general del sistema homogeni associat, cosa que ja sabem des de l'apartat 2.

VII.5 Mètode de Gauss

Un altre mètode per a resoldre sistemes d'equacions és el de reducció o de Gauss-Jordan. La seva justificació teòrica rau en uns raonaments molt simples. Sigui $f : E \longrightarrow F$ una aplicació lineal i siguin $u_1,\ldots,u_n$ i $v_1,\ldots,v_m$ bases de E i F respectivament. Denotem, com sempre, per A la matriu associada a f en aquestes bases i per b el vector de F de coordenades $(b^1,\ldots,b^m)$. Volem trobar les antiimatges x de b : $f(x) = b$. Els canvis en la base de F donen lloc a canvis en la matriu associada i en les coordenades de b. Les coordenades de x resten invariables. Per tant, si A_1 i b_1 són la nova matriu associada i les noves coordenades de b, les solucions del sistema $A_1x = b_1$ coincideixen amb les del sistema original $Ax = b$. Efectuant canvis en la base de F podem aconseguir un sistema amb una matriu prou simple perquè el fet de trobar la solució general no impliqui cap càlcul. A més, els canvis que efectuarem són únicament de tres tipus molt simples:

1. Permutació de l'ordre dels vectors de la base de F. Naturalment, aleshores, les coordenades de
$$f(u_i) = a_i^1 v_1 + \ldots + a_i^m v_m$$
i de
$$b = b^1 v_1 + \ldots + b^m v_m$$
queden permutades, que equival al fet que a (A, b) les files quedin permutades.

2. Substitució d'un vector v_j de la base per kv_j amb $k \neq 0$. Llavors
$$f(u_i) = a_i^1 v_1 + \ldots + (a_i^j k^{-1}) k v_j + \ldots + a_i^m v_m$$
$$b = b^1 v_1 + \ldots + (b^j k^{-1}) k v_j + \ldots + b^m v_m.$$
És a dir, a la matriu (A, b) la fila j queda multiplicada per k^{-1}.

3. Substitució d'un vector v_j de la base de F per $v_j + kv_h (h \neq j)$. Llavors
$$f(u_i) = a_i^1 v_1 + \ldots + a_i^j (v_j + kv_h) + \ldots + (a_i^h - a_i^j k) v_h + \ldots + a_i^m v_m$$
$$b = b^1 v_1 + \ldots + b^j (v_j + kv_h) + \ldots + (b^h - b^j k) v_h + \ldots + b^m v_m.$$
És a dir, a la matriu (A, b), a la fila h se li resta la fila j multiplicada per k.

Fent canvis del tipus 1, 2 i 3 i permutant, si convé, l'ordre de les incògnites, que equival a permutar l'ordre de les n primeres columnes a (A, b), obtenim una matriu de la forma

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & c_{r+1}^1 & \ldots & c_n^1 & d^1 \\ 0 & 1 & \ldots & 0 & c_{r+1}^2 & \ldots & c_n^2 & d^2 \\ \cdot & \cdot & \ldots & \cdot & \ldots & \ldots & \cdot & \cdot \\ 0 & 0 & \ldots & 1 & c_{r+1}^r & \ldots & c_n^r & d^r \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & d^{r+1} \\ \cdot & \cdot & \ldots & \cdot & \ldots & \ldots & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & d^m \end{pmatrix}.$$

El sistema

$$\begin{cases} x^1 + c_{r+1}^1 x^{r+1} + \ldots + c_n^1 x^n = d^1 \\ \ldots\ldots\ldots\ldots \\ x^r + c_{r+1}^r x^{r+1} + \ldots + c_n^r x^n = d^r \\ 0 = d^{r+1} \\ \ldots \\ 0 = d^m \end{cases}$$

té, doncs, les mateixes solucions que l'original, llevat potser de l'ordre de les incògnites. Per tant, el sistema és compatible si i només si
$$d^{r+1} = \ldots = d^m = 0$$

i, en aquest cas, la solució general és

$$x^i = d^i - c^i_{r+1}x^{r+1} - \ldots - c^i_n x^n \quad i = 1, \ldots, r.$$

Nota:

Els canvis en la matriu (A, b) es fan de la manera següent: si la primera columna és tota 0 es passa al lloc n. Si hi ha un element no nul, es permuten les files de forma que quedi en primer lloc. Amb un canvi del tipus 2 es pot aconseguir que aquest element passi a ser un 1 i amb canvis del tipus 3 es pot aconseguir que la resta de la columna sigui 0. La primera columna queda, així, en la forma desitjada. Suposem que tenim h columnes en la forma desitjada. Si a la columna $h+1$ els els elements de les files $h+1, \ldots, m$ són 0, la posem al lloc n. En cas contrari col·loquem un element no nul a la fila $h+1$, *permutant únicament les files* $h+1, \ldots, m$. Amb canvis del tipus 2 i 3 podem aconseguir que aquest element sigui 1 i la resta de la columna sigui 0. Observem que d'aquesta forma les columnes anteriors no varien. El procés pot continuar fins a obtenir una matriu com la que hem escrit més amunt.

Un dels avantatges del mètode de Gauss és que es pot aplicar simultàniament a sistemes d'equacions amb la mateixa matriu i diferents termes independents. Siguin, per exemple, $Ax = b$ i $Ax = c$ dos sistemes amb matriu A. Si fem els canvis necessaris a la matriu (A, b, c) resoldrem a la vegada els dos sistemes.

Exemple:

Considerem el sistema

$$\begin{cases} x^1 - 2x^2 + 3x^3 + 5x^4 - 4x^5 &= b^1 \\ 2x^1 - 4x^2 + 6x^3 + 5x^4 + 2x^5 &= b^2 \\ 2x^1 - 5x^2 + 7x^3 + 7x^4 + 3x^5 &= b^3 \\ -x^1 + x^2 - 2x^3 - 3x^4 + 5x^5 &= b^4 \end{cases}$$

i suposem que ens interessen les solucions quan els termes (b^1, b^2, b^3, b^4) són $(2, -6, -7, -3)$ i $(-3, -1, 1, 2)$. Escrivim:

$$\begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \hline \left(\begin{array}{ccccccc} 1 & -2 & 3 & 5 & -4 & 2 & -3 \\ 2 & -4 & 6 & 5 & 2 & -6 & -1 \\ 2 & -5 & 7 & 7 & 3 & -7 & 1 \\ -1 & 1 & -2 & -3 & 5 & -3 & 2 \end{array}\right). \end{array}$$

Hem escrit una primera fila que indica l'ordre de les columnes corresponents a la matriu A. Comencem, per exemple, restant de la segona, tercera i quarta

fila la primera multiplicada per 2, 2 i -1 respectivament:

$$\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & & \end{array} \\ \left(\begin{array}{ccccccc} 1 & -2 & 3 & 5 & -4 & 2 & -3 \\ 0 & 0 & 0 & -5 & 10 & -10 & 5 \\ 0 & -1 & 1 & -3 & 11 & -11 & 7 \\ 0 & -1 & 1 & 2 & 1 & -1 & -1 \end{array}\right) \end{array}.$$

Per continuar hem de canviar l'ordre, per exemple, de les files segona i tercera. Fent llavors canvis del tipus 2 i 3 obtenim:

$$\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & & \end{array} \\ \left(\begin{array}{ccccccc} 1 & 0 & 1 & 11 & -26 & 24 & -17 \\ 0 & 1 & -1 & 3 & -11 & 11 & -7 \\ 0 & 0 & 0 & -5 & 10 & -10 & 5 \\ 0 & 0 & 0 & 5 & -10 & 10 & -8 \end{array}\right) \end{array}.$$

Per poder continuar amb el mètode general explicat a la nota hem de canviar l'ordre de les columnes. Fixem-nos, però, que si aquí sumem les dues últimes files, obtenim com a última fila

$$(\ 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad -3\ ).$$

Això ens diu que el sistema, amb la segona sèrie de termes independents, és incompatible. Continuem, per tant, només amb la primera columna de termes independents. Suprimim també l'última fila de zeros que correspon a l'"equació" $0 = 0$. Queda

$$\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 4 & 3 & 5 & & \end{array} \\ \left(\begin{array}{ccccccc} 1 & 0 & 11 & 1 & -26 & 24 \\ 0 & 1 & 3 & -1 & -11 & 11 \\ 0 & 0 & -5 & 0 & 10 & -10 \end{array}\right) \end{array},$$

d'on resulta

$$\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 4 & 3 & 5 & & \end{array} \\ \left(\begin{array}{ccccccc} 1 & 0 & 0 & 1 & -4 & 2 \\ 0 & 1 & 0 & -1 & -5 & 5 \\ 0 & 0 & 1 & 0 & -2 & 2 \end{array}\right) \end{array}$$

i, per tant, les solucions del primer sistema donat són les solucions del sistema

$$\left\{\begin{array}{rcl} x^1 + x^3 - 4x^5 & = & 2 \\ x^2 - x^3 - 5x^5 & = & 5 \\ x^4 - 2x^5 & = & 2; \end{array}\right.$$

és a dir,

$$\begin{cases} x^1 &= 2 - x^3 + 4x^5 \\ x^2 &= 5 + x^3 + 5x^5 \\ x^4 &= 2 + 2x^5. \end{cases}$$

VII.6 Càlcul de la matriu inversa

Donada $A \in M_{n\times n}(K)$ es tracta de trobar, si existeix, una matriu $B = (b_i^j)$ que faci $AB = I$. Això equival a buscar les n columnes $b_i = (b_i^1, \ldots, b_i^n)$ de B de forma que

$$Ab_i = e_i$$

on $e_i = (0, \ldots, 1, \ldots, 0)$ és la columna i de I. En altres paraules, hem de resoldre els n sistemes

$$Ax = e_i \qquad i = 1, \ldots, n,$$

tots amb la mateixa matriu. Fem-ho pel mètode de Gauss. Considerem la matriu $(A, e_1, \ldots, e_n)$ i modifiquem-la com a l'apartat anterior. Observem que $(e_1, \ldots, e_n)$ és precisament la matriu identitat. Partim, doncs, de

$$(A, I)$$

i arribarem a una matriu del tipus (I, B):

$$\begin{pmatrix} 1 & \ldots & 0 & b_1^1 & \ldots & b_n^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \ldots & 1 & b_1^n & \ldots & b_n^n \end{pmatrix} \neq 0.$$

La solució del primer sistema, és a dir, la primera columna de B, és $x^i = b_1^i$, $i = 1, \ldots, n$, etc. En resum, resulta que la matriu (b_i^j) que hem obtingut és precisament la matriu inversa buscada. Naturalment, pot passar que la matriu A no es pugui transformar en la matriu identitat I. Aleshores un dels sistemes $Ax = e_i$ resulta incompatible i A no té matriu inversa.

Exercici:

Proveu que si A no té inversa un dels sistemes $Ax = e_i$ és incompatible.

VII.7 Nota històrica

L'estudi dels sistemes d'equacions lineals fou iniciat per Gottfried Wilhelm Leibniz (1646–1716). Ja el 1693 Leibniz considera un sistema de tres equacions lineals amb dues incògnites, elimina les incògnites i obté un determinant (la resultant del sistema).

La solució de sistemes d'equacions lineals utilitzant el que avui anomenemen determinants fou ideada per Colin Maclaurin (1698–1746) el 1729. Gabriel Cramer (1704–1752) primer, i després el 1764 Étienne Bézout (1730–1783), van demostrar que un sistema homogeni quadrat té solució no trivial si i només si el determinant del sistema s'anul·la. També durant aquest segle Jean-Baptiste le Rond d'Alembert (1717–1783) demostra que la solució general d'un sistema s'obté sumant una solució particular a les solucions del sistema homogeni associat.

L'existència i nombre de solucions foren temes discutits per Henry J.S. Smith (1826–1883) el 1861 en termes dels rangs de la matriu del sistema i de la matriu ampliada. La major part de resultats en aquest sentit es deuen a Leopold Kronecker (1823–1891) i a Arthur Cayley (1821–1895) i apareixen ja el 1867 en el llibre de Charles L. Dodgson (Lewis Carroll (1832-1898), l'autor d'*Alícia al país de les meravelles*) *An elementary theory of determinants.*

VII.8 Exercicis

1. Donat el sistema

$$\left\{\begin{array}{rcl} x+by+az & = & 1 \\ ax+by+z & = & a \\ x+aby+z & = & b. \end{array}\right.$$

a) Per a quins valors de a i b el sistema té solució?

b) Resoleu-lo i determineu quan té una única solució.

2. Discutiu el sistema homogeni

$$\left\{\begin{array}{rcl} -6x-6y+(11-a)z & = & 0 \\ 3x+(12-a)y-6z & = & 0 \\ (2-a)x+3y-6z & = & 0 \end{array}\right.$$

i trobeu-ne les solucions.

3. Resoleu els sistemes de congruències

a)
$$\left.\begin{array}{r} x+2y+z \equiv 1 \\ 2x+y+2z \equiv 1 \\ y+2z \equiv 1 \end{array}\right\} \text{ mòd } 5$$

b)
$$\left.\begin{array}{r} x+2y+z \equiv -1 \\ 2x+y+2z \equiv 1 \\ y+2z \equiv 1 \end{array}\right\} \text{ mòd } 3.$$

4. Trobeu un sistema d'equacions lineals homogeni, les solucions del qual siguin exactament els vectors del subespai vectorial de $(\mathbf{Z}/(7))^4$ generat per

$$(1,0,1,-1), \quad (2,1,3,0) \quad \text{i} \quad (1,3,4,5).$$

5. Discutiu segons els valors del paràmetre a el sistema de congruències mòdul 5

$$\left.\begin{array}{r} x+y+3z \equiv 2 \\ 2x+3y+4z \equiv 0 \\ 3x+4y+az \equiv 3 \end{array}\right\}.$$

6. Resoleu el sistema d'equacions lineals complexes

$$\left\{\begin{array}{rcl} x+y+iz+t & = & 0 \\ 2x-y+2z-t & = & 1 \\ x+iy-z+it & = & 2 \\ x+y+z-t & = & 0. \end{array}\right.$$

7. Determineu $a \in \mathbf{R}$ per tal que l'endomorfisme f de $\mathbf{R}^3$ definit per

$$f(x,y,z) = (ax+y+z,\ x+ay+z,\ x+y+az)$$

tingui nucli de la màxima dimensió possible, i doneu-ne una base.

8. Sigui $A \in M_{n\times n}(K)$ una matriu de rang $r < n$. Demostreu que existeix una matriu $B \in M_{n\times n}(K)$, com a mínim, $B \neq 0$, tal que $AB = 0$. Quin és el màxim rang d'una tal matriu B?

9. Sigui $A \in M_{n\times n}(K)$ una matriu de rang $n-1$. Demostreu que, escollides dues files qualssevol de A, les n-ples formades pels adjunts dels seus elements són proporcionals.

10. Discutiu i resoleu el sistema

$$\left\{\begin{array}{rcll} ax^i + a_i x^{n+1} & = & b_i & i = 1,2,\ldots,n \\ a_1x^1 + \ldots + a_nx^n + ax^{n+1} & = & b_{n+1} & \end{array}\right.$$

per als diferents valors de a, a_i, b_j.

VII.9 Exercicis de programar

11. Escriviu un programa que doni la solució de sistemes 2×2 o bé 3×3 (reals) amb determinant diferent de zero. (Indicació: comenceu calculant el determinant. Si és 0, atureu el programa; si no, apliqueu la regla de Cramer de la secció 3).

Nota:

Per a n gran, aquest mètode és molt ineficient.

12. (Mètode de Gauss). Aquest programa ha de servir per a

a) Donar la solució general d'un sistema d'equacions.

b) Calcular el rang d'una matriu (VI.5).

c) Invertir una matriu (VII.6).

d) Trobar una base de l'espai de solucions d'un sistema homogeni.

Preparació:

- Entreu una matriu real A no necessàriament quadrada.
- Entreu una família de vectors $b_1, \ldots, b_k$ de $\mathbf{R}^n$.
- Construïu la matriu ampliada $(A, b_1, \ldots, b_k)$. Efectueu-hi els canvis 1, 2, 3 del § 5 per reduir A a una matriu de la forma corresponent.
- Apliqueu aquesta situació a cada un dels tres primers temes esmentats.
- Si el sistema és homogeni ($i = 1$, $b_1 = \vec{0}$), llavors la dimensió de l'espai de solucions és $m-r$. Una base d'aquest espai s'obté donant successivament el valor 1 a cada una de les incògnites lliures i 0 a les restants.

Notes:

a) Convé preparar aquest programa de manera que pugui ésser utilitzat dins d'altres programes sempre que convingui.

b) Si anem guardant els canvis de signe i els escalars eliminats, podem utilitzar aquest programa per calcular $\det A$.

c) Els elements que van quedant a la diagonal de A s'anomenen "pivots". Per tal de minimitzar la propagació d'errors d'arrodoniment, és convenient efectuar a cada etapa permutacions de files de manera que quedi com a pivot l'element de valor absolut més gran entre tots els disponibles a la columna corresponent.

Capítol VIII

Estructura dels endomorfismes

Donada una aplicació lineal $f : E \longrightarrow F$ podem sempre escollir bases de E i F en les quals la matriu de f sigui extraordinariament simple. En efecte, sigui $u_1, \ldots, u_k$ una base de $\operatorname{Nuc} f$ i $u_1, \ldots, u_k, u_{k+1}, \ldots, u_n$ una base de E. Llavors, (V.1.1), $f(u_{k+1}), \ldots, f(u_n)$ són vectors de F linealment independents. Completem-los a una base de F: $f(u_{k+1}), \ldots, f(u_n), v_{n-k+1}, \ldots, v_m$. La matriu de f en aquestes bases és:

$$\begin{pmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 1 \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

En estudiar endomorfismes $f : E \longrightarrow E$ és natural, però, exigir que els vectors $u \in E$ i llurs imatges estiguin expressats en la mateixa base. Aleshores no es pot aconseguir, en general, una matriu tan simple com la que acabem de trobar. En tot aquest capítol, E denotarà un espai vectorial de dimensió n sobre un cos K.

VIII.1 Vectors propis i valors propis. Polinomi característic

Sigui $f \in \operatorname{End}(E)$. Un vector $v \in E$, $v \neq \vec{0}$, és un *vector propi* de f si

$$f(v) = kv, \quad k \in K.$$

Direm, aleshores, que k és un *valor propi* de f.

Exemples:

1. Els vectors de $\operatorname{Nuc} f$ diferents de $\vec{0}$ són vectors propis de valor propi 0.
2. Si $f = kI$ (homotècia de raó k), tot $v \neq \vec{0}$ és un vector propi de f, i k és l'únic valor propi de f.

Exercici:

Si tot $v \in E$, $v \neq \vec{0}$, és vector propi de f, f és una homotècia.

Un vector $v \neq \vec{0}$ és vector propi de f de valor propi k si i només si $f(v) - kv = \vec{0}$, és a dir, si i només si $v \in \operatorname{Nuc}(f - kI)$. Un element $k \in K$ és un valor propi si i només si $\operatorname{Nuc}(f - kI) \neq \{\vec{0}\}$. Es diu *multiplicitat* del valor propi k la dimensió de $\operatorname{Nuc}(f - kI)$.

Proposició 1.1 *$k \in K$ és valor propi de f si i només si $\det(f - kI) = 0$.*

DEMOSTRACIÓ: k és valor propi $\Leftrightarrow \operatorname{Nuc}(f - kI) \neq \{\vec{0}\} \Leftrightarrow \det(f - kI) = 0$ per (VI.3.4). □

Sigui $A = (a_i^j)$ la matriu de f en una certa base $e_1, \ldots, e_n$ de E. Llavors,

$$\det(f - kI) = \begin{vmatrix} a_1^1 - k & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 - k & \ldots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \ldots & a_n^n - k \end{vmatrix} = 0.$$

Aquesta expressió és una equació de grau n en la incògnita k, el costat esquerre de la qual és el valor a k d'un polinomi $p_A(x)$, que anomenarem *polinomi característic* de A. Si B és la matriu associada a f en una altra base, veurem que $p_B(x) = p_A(x)$. Això ens permetrà parlar del *polinomi característic de f*, $p_f(x)$. Tenim

$$\det(B - kI) = det(f - kI) = det(A - kI) \qquad \forall k \in K.$$

És a dir, $p_B(k) = p_A(k)$, $\forall k \in K$. Si K té més de n elements, resulta que $p_B(x) = p_A(x)$ (II.5.3). Si K té menys de n elements aquesta demostració no serveix, però el resultat continua essent cert. Una manera de provar-ho és considerant $p_A(x)$ com un determinant amb elements a l'anell $K[x]$:

$$p_A(x) = \det(A - xI) = \begin{vmatrix} a_1^1 - x & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 - x & \ldots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \ldots & a_n^n - x \end{vmatrix}.$$

La definició i la majoria de les propietats dels determinants de matrius amb elements en un cos K es poden generalitzar a matrius amb elements a $K[x]$. En

particular, es compleix que el determinant d'un producte de matrius és el producte de llurs determinants. Llavors, si A i B són matrius del mateix endomorfisme, $B = P^{-1}AP$ (on P és una matriu invertible) i

$$\begin{aligned}\det(B - xI) &= \det(P^{-1}AP - xI) = \\ &= \det(P^{-1}(A - xI)P) = \det P^{-1} \cdot \det(A - xI) \cdot \det P = \det(A - xI).\end{aligned}$$

Proposició 1.2 *El polinomi característic de A és*

$$p_A(x) = (-1)^n x^n + (-1)^{n-1}(a_1^1 + \cdots + a_n^n)x^{n-1} + \cdots + (-1)^r A_r x^r + \ldots + \det A$$

on A_r és la suma dels determinants dels menors d'ordre $n-r$ formats pels elements de A de $(n-r)$ files i $(n-r)$ columnes corresponents als mateixos índexs: a_i^j $i = i_1, \ldots, i_{n-r}$, $j = i_1, \ldots, i_{n-r}$. És a dir, són els menors d'ordre $n-r$ que tenen la diagonal principal sobre la diagonal principal de A.

No farem el càlcul dels A_r, que és llarg i pesat. A més a més, existeixen formes més còmodes de trobar-los que les que podríem donar amb els coneixements que tenim ara. Ens limitarem a calcular el coeficient de x^{n-1} i el terme independent.

De la definició de determinant resulta

$$\begin{vmatrix} a_1^1 - k & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 - k & \ldots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \ldots & a_n^n - k \end{vmatrix} = (a_1^1 - k)(a_2^2 - k)\cdots(a_n^n - k) + S$$

on S és una suma de productes a cada un dels quals hi ha $n-2$ o menys elements de la diagonal. Els termes de grau n i $n-1$ en k són, per tant,

$$(-1)^n k^n + (-1)^{n-1}(a_1^1 + \cdots + a_n^n)k^{n-1}.$$

El terme independent de $p_A(x)$ és $p_A(0) = \det(A - 0I) = \det A$.

Dues matrius A i B associades al mateix endomorfisme, és a dir, tals que $B = P^{-1}AP$ amb P invertible, es diuen *equivalents*. Aleshores, de (1.2) i de la invariància del polinomi característic, es dedueix que $A_r = B_r$ i, en particular,

$$a_1^1 + \cdots + a_n^n = b_1^1 + \ldots + b_n^n.$$

La suma $a_1^1 + \cdots + a_n^n$ es diu la *traça de A*, $\operatorname{tr} A$. Com que dues matrius equivalents tenen la mateix traça, té sentit parlar de la *traça de f*, $\operatorname{tr} f$.

VIII.2 Diagonalització de matrius

Sigui $f \in \mathrm{End}(E)$. Si aconseguim una base de E amb vectors propis de f, la matriu de f tindrà una forma molt simple

$$\begin{pmatrix} k_1 & 0 & \dots & 0 \\ 0 & k_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & k_n \end{pmatrix}.$$

Tots els elements no situats sobre la diagonal seran 0 i $k_1, k_2, \dots, k_n$ seran els valors propis dels vectors propis de la base.

Una matriu d'aquest tipus es diu una matriu *diagonal*. Diagonalitzar f vol dir trobar una base de vectors propis de f; diagonalitzar una matriu A vol dir trobar una matriu diagonal equivalent a A. Direm que un endomorfisme f és *diagonalitzable* si es pot trobar una base de vectors propis de f. Una matriu es pot diagonalitzar si i només si l'endomorfisme associat és diagonalitzable.

A l'apartat anterior hem donat ja una manera de trobar els valors propis i els vectors propis: els valors propis són els zeros del polinomi característic (1.1) i $\mathrm{Nuc}(f - kI)$ és el conjunt de vectors propis de valor propi k (més el $\vec{0}$). Només resta, doncs, veure si hi ha n vectors propis linealment independents.

Proposició 2.1 *Vectors propis de valors propis diferents són linealment independents.*

DEMOSTRACIÓ: Siguin $v_1, \dots, v_m$ vectors propis de valors propis $k_1, \dots, k_m$ diferents. Procedirem per inducció sobre m. Si $m = 1$, $v_1 \neq \vec{0}$ és linealment independent. En general sigui $a^1 v_1 + \cdots + a^m v_m = \vec{0}$; aleshores

$$\vec{0} = (f - k_1 I)(a^1 v_1 + \cdots + a^m v_m) = a^2(k_2 - k_1)v_2 + \cdots + a^m(k_m - k_1)v_m.$$

Per hipòtesi d'inducció $v_2, \dots, v_m$ són linealment independents, d'on $a^j(k_j - k_1) = 0$, $j = 2, \dots, m$. Com que $k_j \neq k_1$ si $j \neq 1$, cal que

$$a^j = 0 \qquad \text{per a } j = 2, \dots, m.$$

Llavors $a^1 v_1 = \vec{0}$, d'on també $a^1 = 0$. □

Corol·lari 2.2 *El nombre de valors propis diferents és $\leq n$. Si hi ha exactament n valors propis diferents, l'endomorfisme és diagonalitzable.*□

Exemple:

1. Considerem l'endomorfisme $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ amb matriu

$$\begin{pmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{pmatrix}$$

en la base usual $(1,0)$, $(0,1)$. El seu polinomi característic és

$$x^2 - 1 = (x-1)(x+1).$$

f té, per tant, dos valors propis: $+1$ i -1. El subespai de vectors propis de valor propi $+1$ és $\mathrm{Nuc}(f - I)$. La matriu de $(f - I)$ és

$$\begin{pmatrix} -2/5 & 4/5 \\ 4/5 & -8/5 \end{pmatrix}.$$

D'on resulta que $\mathrm{Nuc}(f - I) = \{(2y, y)\} = \langle (2,1) \rangle$. Anàlogament, es veu que $\mathrm{Nuc}(f+I) = \langle (-1,2) \rangle$ és el subespai de vectors propis de valor propi -1. Els vectors $(2,1)$, $(-1,2)$ formen una base en la qual la matriu de f és

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

La imatge d'un $v \in E$ es pot trobar geomètricament de la forma següent: descomponem v com a suma d'un vector v_1 de $\langle (2,1) \rangle$ i un vector v_2 de $\langle (-1,2) \rangle$: $v = v_1 + v_2$. Llavors $f(v) = v_1 - v_2$. Això és una simetria d'eix $\langle (2,1) \rangle$.

2. Considerem ara $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ amb matriu

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

en la base usual $(1,0)$,$(0,1)$. El polinomi característic és $(1-x)^2$ i, per tant, l'únic valor propi és 1. Si f diagonalitzés, la seva matriu diagonal seria

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

f seria la identitat i això no és cert. De fet, el subespai de vectors propis és de dimensió 1: $\mathrm{Nuc}(f - I) = \langle (1,0) \rangle$.

Proposició 2.3 *Si r és la multiplicitat del valor propi k, és a dir, si es té $r = \dim \mathrm{Nuc}(f - kI)$, i s és la multiplicitat del zero k del polinomi característic, aleshores, $r \leq s$.*

DEMOSTRACIÓ: Sigui $v_1, \ldots, v_r$ una base de $Nuc(f - kI)$. Completem-la fins a obtenir una base de E: $v_1, \ldots, v_n$. En aquesta base la matriu de f és de la forma

$$A = \begin{pmatrix} k & 0 & \ldots & 0 & a^1_{r+1} & \ldots & a^1_n \\ 0 & k & \ldots & 0 & \vdots & & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & k & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 & a^n_{r+1} & \ldots & a^n_n \end{pmatrix}.$$

El polinomi característic és, doncs, $p(x) = (k-x)^r \cdot q(x)$, que demostra l'enunciat.□

Suposem ara que f és diagonalitzable i sigui

$$\begin{pmatrix} k_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & k_n \end{pmatrix}$$

la seva matriu diagonal ($k_1, \dots, k_n$ no necessàriament diferents). El polinomi característic és

$$p(x) = (k_1 - x) \cdots (k_n - x)$$

i descompon, per tant, en factors lineals. Si un valor propi k_i surt s vegades a la diagonal, la multiplicitat del zero k_i de $p(x)$ és s. D'altra banda, $(f - k_i I)$ tindrà una matriu diagonal amb exactament s zeros a la diagonal; d'on $\dim \mathrm{Nuc}(f - k_i I) = s$. Aquests fets caracteritzen els endomorfismes diagonalitzables, com ho demostra el següent teorema.

Teorema 2.4 (de diagonalització) *Un endomorfisme f és diagonalitzable si i només si el seu polinomi característic descompon en factors lineals i la multiplicitat de cada un dels seus zeros coincideix amb la seva multiplicitat com a valor propi de f.*

DEMOSTRACIÓ: Hem vist ja que aquestes condicions són necessàries. Demostrem ara que són suficients perquè f sigui diagonalitzable. Sigui

$$p(x) = (-1)^n (x - k_1)^{n_1} \cdots (x - k_r)^{n_r} \quad n_1 + \dots + n_r = n$$

el polinomi característic de f. Posem $E_{k_i} = \mathrm{Nuc}(f - k_i I)$. Anem a veure que

$$E = E_{k_1} \oplus \cdots \oplus E_{k_r}.$$

Un cop vist això, podrem obtenir una base de vectors propis agafant bases a $E_{k_1}, \dots, E_{k_r}$. En aquesta base la matriu de f serà diagonal. Provem, doncs, que l'expressió dels vectors de E com a suma de vectors dels E_{k_i} és única (IV.4). En efecte,

$$v_1 + \dots + v_r = w_1 + \dots + w_r$$

amb $v_i, w_i \in E_{k_i}$, $i = 1, \dots, r$, implica $(v_1 - w_1) + \cdots + (v_r - w_r) = \vec{0}$. Els vectors $v_i - w_i$ són vectors propis de valors propis diferents, o $\vec{0}$. Per (2.1) han d'ésser tots zero: $v_i - w_i = \vec{0}$; és a dir, $v_i = w_i$, $i = 1, \dots, r$. La suma dels E_{k_i} és, doncs, directa i la seva dimensió $n_1 + \dots + n_r = n$. Per tant,

$$E_{k_1} \oplus \cdots \oplus E_{k_r} = E.\square$$

Una matriu $A = (a_i^j)$ es diu *triangular superior* si $a_i^j = 0$, per a $i < j$. Direm que un endomorfisme és *triangulable* si, en una base convenient, la seva matriu és triangular.

Teorema 2.5 (de triangulació) *Un endomorfisme és triangulable si i només si el seu polinomi característic descompon en factors de primer grau.*

DEMOSTRACIÓ: Si l'endomorfisme f té una matriu triangular

$$\begin{pmatrix} a_1^1 & a_2^1 & a_3^1 & \dots & a_n^1 \\ 0 & a_2^2 & a_3^2 & \dots & a_n^2 \\ 0 & 0 & a_3^3 & \dots & a_n^3 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_n^n \end{pmatrix},$$

el seu polinomi característic, $p(x) = (a_1^1 - x)(a_2^2 - x)\cdots(a_n^n - x)$, descompon en factors lineals. Provarem el recíproc per inducció sobre n. Per a $n = 1$, tota matriu és triangular. Sigui $n \geq 2$ qualsevol. El polinomi característic té com a mínim una arrel; hi ha doncs, com a mínim, un valor propi. Sigui v_1 un vector propi, i $v_1, v_2, \dots, v_n$ una base de E. La matriu de f en aquesta base és del tipus

$$\begin{pmatrix} k & a_2^1 & \dots & a_n^1 \\ 0 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_2^n & \dots & a_n^n \end{pmatrix}.$$

Considerem ara l'aplicació

$$g : \langle v_2, \dots, v_n \rangle \longrightarrow \langle v_2, \dots, v_n \rangle$$

de matriu

$$\begin{pmatrix} a_2^2 & \dots & a_n^2 \\ \vdots & \ddots & \vdots \\ a_2^n & \dots & a_n^n \end{pmatrix}.$$

Tenim que $\det(f - xI) = (k-x)\cdot\det(g - xI)$, i, per tant, el polinomi característic de g, $\det(g - xI)$, descompon també en factors de primer grau. Per hipòtesi d'inducció existeix llavors una base $u_2, \cdots, u_n$ de $\langle v_2, \cdots, v_n \rangle$ en la qual la matriu de g és triangular:

$$\begin{pmatrix} b_2^2 & b_3^2 & \dots & b_n^2 \\ 0 & b_3^3 & \dots & b_n^3 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_n^n \end{pmatrix}.$$

Ara bé: $f(v_i) = a_i^1 v_1 + g(v_i)$, $i = 2, \dots, n$. Per tant, si $u_j = \sum_{i=2}^n c_j^i v_i$,

$$f(u_j) = \sum_{i=2}^n c_j^i f(v_i) = \sum_{i=2}^n c_j^i (a_i^1 v_1 + g(v_i)) = \left(\sum_{i=2}^n c_j^i a_i^1 \right) v_1 + g(u_j),$$

amb $j = 2, \ldots, n$. La matriu de f en la base $v_1, u_2, \ldots, u_n$ s'obté, doncs, afegint a la matriu de g una primera columna $(k, 0, \ldots, 0)$ i una primera fila $(k, b_2^1, \ldots, b_n^1)$ on $b_j^1 = \sum_{i=2}^n c_j^i a_i^1$, $j \geq 2$. És, per tant, una matriu triangular. □

Corol·lari 2.6 *Tot endomorfisme d'un espai vectorial sobre els complexos és triangulable.*□

VIII.3 Polinomi mínim

L'estudi dels vectors propis ens ha permès simplificar la matriu d'un endomorfisme en molts casos. Queda, però, sense resoldre el cas general. Els nostres passos s'encaminen ara cap a l'obtenció d'uns teoremes de descomposició de E en suma directa de subespais que permetin obtenir bases convenients on es pugui expressar l'endomorfisme. Observem que una descomposició d'aquest tipus és la que ens ha permès demostrar el teorema de diagonalització (2.4).

Considerem les potències de f: f^r, $r = 0, 1, 2 \ldots$,

$$f^0 = I,\ f^1 = f,\ \ldots,\ f^r = f \circ f^{r-1}.$$

Si la dimensió de E és n, l'espai vectorial $\mathrm{End}(E)$ és de dimensió n^2, i les potències f^r no poden ésser totes linealment independents. Les combinacions lineals

$$a_0 I + a_1 f + \ldots + a_s f^s = 0$$

ens porten a considerar el nucli de l'aplicació

$$\begin{array}{rcl} \Phi_f : K[x] & \longrightarrow & \mathrm{End}(E) \\ p(x) = a_0 + a_1 x + \ldots + a_r x^r & \longmapsto & p(f) = a_0 I + a_1 f + \ldots + a_r f^r. \end{array}$$

Es compleixen les següents propietats:

- $\Phi_f(p(x) + q(x)) = p(f) + q(f) = \Phi_f(p(x)) + \Phi_f(q(x))$.
- $\Phi_f(p(x) \cdot q(x)) = p(f) \circ q(f) = \Phi_f(p(x)) \circ \Phi_f(q(x))$.
- $\Phi_f(kp(x)) = kp(f) = k\Phi_f(p(x))$.

Φ_f és, doncs, un morfisme d'àlgebres (V.5). De la commutativitat del producte de $K[x]$ es dedueix que dos endomorfismes de la imatge sempre commuten:

$$p(f) \circ g(f) = g(f) \circ p(f).$$

El nucli de Φ_f és un ideal de $K[x]$ i per (II.2.2)

$$\mathrm{Nuc}\,\Phi_f = \{p(x) \in K[x] \mid p(f) = 0\} = (m_f(x)).$$

Els polinomis de $\mathrm{Nuc}\,\Phi_f$ es diuen *polinomis anul·ladors de* f; $m_f(x)$ es diu el *polinomi mínim de* f i està determinat llevat de factors de K (II.2). En general s'agafa $m_f(x)$ mònic, és a dir, amb coeficient de màxim grau igual a 1.

Exemples:

1. Si $E = \{\vec{0}\}$ i f és l'únic endomorfisme de $\{\vec{0}\}$, $\operatorname{Nuc}\Phi_f = K[x] = (a_0)$, $a_0 \neq 0$.
   Recíprocament, si el polinomi mínim de f és $a_0 \in K$, $\operatorname{Nuc}\Phi_f = (a_0) = K[x]$ i, en particular, $0 = \Phi_f(1) = I_E$. Això implica que $E = \{\vec{0}\}$.
2. Si $E \neq \{\vec{0}\}$ i $f = 0$, $\operatorname{Nuc}\Phi_f = (x)$.
3. Si $E \neq \{\vec{0}\}$ i $f = I$, $x - 1 \in \operatorname{Nuc}\Phi_f$ i $m_f(x) \mid (x-1)$. Però, com que $m_f(x)$ no és constant, cal que $m_f(x) = x - 1$.
4. Si $E \neq \{\vec{0}\}$, $f = kI$ si i només si $m_f(x) = x - k$.

Fixat un vector $u \in E$ considerem ara l'aplicació

$$\begin{array}{rccc} \Phi_u : & K[x] & \longrightarrow & E \\ & p(x) & \longmapsto & p(f)(u). \end{array}$$

El nucli de Φ_u és un ideal de $K[x]$

$$\operatorname{Nuc}\phi_u = \{p(x) \in K[x] \mid p(f)(u) = \vec{0}\} = (m_u(x)).$$

$m_u(x)$ es diu el *polinomi mínim de f a u* o simplement el *polinomi mínim de u* (si no hi ha confusió sobre què és f); està determinat llevat de factors de K i generalment s'agafa mònic.

Proposició 3.1 *Sigui*

$$m_u(x) = a_0 + a_1 x + \ldots + a_s x^s$$

el polinomi mínim de u. Aleshores $u, f(u), \ldots, f^{s-1}(u)$ són linealment independents i $u, f(u), \ldots, f^{s-1}(u), f^t(u)$ $(t \geq s)$ són linealment dependents.

DEMOSTRACIÓ: Si $u, f(u), \ldots, f^{s-1}(u)$ fossin linealment dependents hi hauria un polinomi $p(x)$, de grau $< s$, tal que $p(f)(u) = 0$. Això contradiu la definició de $m_u(x)$.

Si $t = s$, $m_u(f)(u) = a_0 u + a_1 f(u) + \ldots + a_s f^s(u) = 0$ ens diu que aquests vectors són linealment dependents. Per a $t > s$ procedirem per inducció. Aleshores,

$$f^{t-1}(u) \in \langle u, f(u), \ldots, f^{s-1}(u)\rangle,$$

d'on

$$f^t(u) \in \langle f(u), f^2(u), \ldots, f^s(u)\rangle = \langle u, f(u), \ldots, f^{s-1}(u)\rangle$$

i $u, f(u), \ldots, f^{s-1}(u), f^t(u)$ són linealment dependents. □

VIII.4 Subespais invariants

Sigui $f \in \mathrm{End}(E)$. Un subespai F de E es diu *invariant per* f si $f(F) \subset F$. En aquest cas f indueix un endomorfisme de F

$$\begin{aligned} f' = f \mid F : F &\longrightarrow F \\ v &\longmapsto f(v) \end{aligned}$$

que anomenarem *restricció de* f *a* F.

Proposició 4.1 *$m_{f'}(x)$ divideix $m_f(x)$.*

DEMOSTRACIÓ: $m_f(f) = 0 \Rightarrow m_f(f)(u) = \vec{0}\ \forall u \in E \Rightarrow m_f(f')(v) = m_f(f)(v) = \vec{0}\ \forall v \in F \Rightarrow m_f(x) \in (m_{f'}(x))$. □

Corol·lari 4.2 *Si dos subespais F i G de E, invariants per f, tenen polinomis mínims primers entre ells, $F \cap G = \{\vec{0}\}$.*

DEMOSTRACIÓ: Clarament $F \cap G$ és també invariant i per (4.1) el seu polinomi mínim és 1. Als exemples del §3 vam veure que, aleshores, l'espai ha d'ésser $\{\vec{0}\}$.□

Proposició 4.3 *Per a tot polinomi $p(x) \in K[x]$, els subespais* $\mathrm{Nuc}\, p(f)$ *i* $\mathrm{Im}\, p(f)$ *són invariants per f.*

DEMOSTRACIÓ: Si $u \in \mathrm{Nuc}\, p(f)$, $p(f)(f(u)) = f(p(f)(u)) = f(\vec{0}) = \vec{0}$, d'on $f(u) \in \mathrm{Nuc}\, p(f)$. Si $u = p(f)(v) \in \mathrm{Im}\, p(f)$, $f(u) = f(p(f)(v)) = p(f)(f(v))$, d'on $f(u) \in \mathrm{Im}\, p(f)$. □

Suposem ara que el polinomi mínim de f, $m_f(x)$, descompon en producte de dos factors primers entre ells:

$$m_f(x) = p(x) \cdot q(x).$$

Considerem els subespais invariants $\mathrm{Nuc}(p(f))$ i $\mathrm{Nuc}(q(f))$. Els polinomis $p(x)$ i $q(x)$ són anul·ladors de la restricció de f a aquests subespais. Per tant, (4.2),

$$\mathrm{Nuc}\, p(f) \cap \mathrm{Nuc}\, q(f) = \{\vec{0}\}.$$

Ara bé, és fàcil veure que $\mathrm{Nuc}\, p(f) \supset \mathrm{Im}\, q(f)$ i que $\mathrm{Nuc}\, q(f) \supset \mathrm{Im}\, p(f)$. Comprovem la primera inclusió: si $u = q(f)(v) \in \mathrm{Im}\, q(f)$, aleshores $p(f)(u) = p(f)q(f)(v) = m_f(f)(v) = \vec{0}$ i, per tant, $u \in \mathrm{Nuc}\, p(f)$. Aquestes inclusions indiquen que

$$\begin{aligned} n &= \dim \mathrm{Nuc}\, p(f) + \dim \mathrm{Im}\, p(f) \\ &\leq \dim \mathrm{Nuc}\, p(f) + \dim \mathrm{Nuc}\, q(f) = \dim(\mathrm{Nuc}\, p(f) \oplus \mathrm{Nuc}\, q(f)) \leq n. \end{aligned}$$

Les dues inclusions són, doncs, igualtats i

$$E = \mathrm{Nuc}\, p(f) \oplus \mathrm{Nuc}\, q(f).$$

Quins són els polinomis mínims de la restricció de f a aquests dos subespais invariants en què es descompon E? Ja hem dit abans que han d'ésser divisors de $p(x)$ i de $q(x)$ (aquests polinomis són anul·ladors): siguin $\bar{p}(x)$ i $\bar{q}(x)$. Però, llavors, $\bar{p}(x)\cdot\bar{q}(x)$ és un anul·lador de f: si $u \in E$, $u = u_1 + u_2$ amb $u_1 \in \operatorname{Nuc} p(f)$, $u_2 \in \operatorname{Nuc} q(f)$, i, aleshores,

$$\bar{p}(f)\bar{q}(f)(u) = \bar{q}(f)\bar{p}(f)(u_1) + \bar{p}(f)\bar{q}(f)(u_2) = \vec{0} + \vec{0} = \vec{0}.$$

Per tant, per una banda $\bar{p}(x)\cdot\bar{q}(x) \in (m_f(x))$ i per l'altra divideix $m_f(x) = p(x)\cdot q(x)$. cal, doncs, que $\bar{p}(x)\cdot\bar{q}(x) = m_f(x)$; és a dir, $\bar{p}(x) = p(x)$ i $\bar{q}(x) = q(x)$ són els polinomis mínims buscats.

Naturalment, si ara $p(x)$ (o $q(x)$) descompon en factors primers, podem descompondre $\operatorname{Nuc} p(f)$ (o $\operatorname{Nuc} q(f)$) en suma de subespais invariants i procedir així tantes vegades com puguem. Tenim d'aquesta forma el

Teorema 4.4 (primer teorema de descomposició) *Si el polinomi mínim de* $f \in \operatorname{End}(E)$ *és*

$$m_f(x) = m_1(x)^{n_1}\cdots m_r(x)^{n_r},$$

on $m_1(x),\ldots,m_r(x)$ *són factors irreductibles, l'espai* E *és suma directa de subespais invariants*

$$E = E^1 \oplus \cdots \oplus E^r,$$

de forma que el polinomi mínim de la restricció de f *a* E^i *és* $m_i(x)^{n_i}$. *Aquesta descomposició és única:*

$$E^i = \operatorname{Nuc}(m_i(f)^{n_i}) \qquad i = 1,\ldots,r.$$

DEMOSTRACIÓ: L'únic que resta per demostrar és la unicitat de la descomposició. Suposem, doncs, que tenim donada una descomposició en subespais invariants

$$E = E^1 \oplus \cdots \oplus E^r$$

de la qual només sabem que el polinomi mínim de la restricció de f a E^i és $m_i(x)^{n_i}$, per a $i = 1,\ldots,r$. Aquesta darrera condició implica que $E^i \subset \operatorname{Nuc}(m_i(f)^{n_i})$; d'on

$$\begin{aligned} n &= \dim E^1 + \ldots + \dim E^r \\ &\leq \dim \operatorname{Nuc}(m_1(f)^{n_1}) + \ldots + \dim \operatorname{Nuc}(m_r(f)^{n_r}) = n. \end{aligned}$$

La desigualtat ha d'ésser, doncs, una igualtat i totes les inclusions anteriors han d'ésser igualtats:

$$E^i = \operatorname{Nuc}(m_i(f)^{n_i}) \qquad i = 1,\ldots,r.\square$$

La descomposició de E en suma directa de subespais invariants redueix l'estudi del comportament de f a l'estudi de les seves restriccions a cada un dels subespais.

Si escrivim la matriu de f en una base de E formada per bases de cada un dels subespais, obtenim

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & 0 \\ & & \ddots & \\ & 0 & & A_r \end{pmatrix}.$$

La matriu A està formada per unes matrius $A_1, \ldots, A_r$ amb la diagonal sobre la de A, i 0 a la resta de llocs. L'estudi de A es redueix, doncs, al de les matrius A_i, que són precisament les matrius de les restriccions de f a cada un dels subespais en què es descompon E.

Aplicarem ara (4.4) a diferents exemples concrets.

Exemples:

1. Considerem l'endomorfisme $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ amb matriu

$$A = \begin{pmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{pmatrix}$$

en la base $(1,0),(0,1)$. (Veure §2). L'estudi de les combinacions lineals entre les seves potències A^n resulta aquí trivial. Tenim $A^2 = I$ i, per tant, $x^2 - 1$ és un polinomi anul·lador. Si $m_f(x) = x - 1$, $m_f(f) = f - I = 0$, d'on $f = I$. Si $m_f(x) = x + 1$, $m_f(f) = f + I = 0$, d'on $f = -I$. Cap dels dos casos no és el nostre. Per tant,

$$m_f(x) = (x-1)(x+1) \qquad \text{i} \qquad E = E^1 \oplus E^2,$$

on $E^1 = \mathrm{Nuc}(f - I) = \langle(2,1)\rangle$, $E^2 = \mathrm{Nuc}(f + I) = \langle(1,-2)\rangle$. La restricció de f a E^1 és I_{E^1}; la restricció a E^2 és $-I_{E^2}$. És a dir, E^1 i E^2 són precisament els subespais dels vectors propis de valors propis 1 i -1 que vam obtenir al §2. Com allà, en la base $(2,1),(1,-2)$ la matriu resulta ésser

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

2. Com a generalització de l'exemple anterior, considerem un endomorfisme f tal que $f^2 = I$ (es diu una *involució*). $m_f(x)$ divideix $x^2 - 1 = (x-1)(x+1)$ i, com a l'exemple 1,

si $m_f(x) = x - 1$, $f = I$;
si $m_f(x) = x + 1$, $f = -I$;
si $m_f(x) = (x-1)(x+1)$, $E = E^1 \oplus E^2$.

El polinomi mínim de la restricció de f a E^1 és $x - 1$ i, per tant, sobre E^1 f és I_{E^1}. Anàlogament, sobre E^2 f és $-I_{E^2}$. Si $e_1, \ldots, e_r$ és una

base de E^1 i $e_{r+1}, \ldots, e_n$ una base de E^2, la matriu de E en la base que resulta de la unió d'aquestes dues és

$$\left(\begin{array}{cccccc} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & -1 & & \\ & 0 & & & \ddots & \\ & & & & & -1 \end{array}\right) \quad \begin{array}{l} \Big\} r \text{ files} \\ \Big\} n-r \text{ files.} \end{array}$$

Observem que aquí també E^1, E^2 són els subespais de vectors propis de valors propis 1, -1 i que la matriu obtinguda és una matriu diagonal.

3. Modifiquem l'exemple anterior estudiant endomorfismes f tals que $f^2 = -I$. El polinomi mínim ara és divisor de x^2+1. Si el cos sobre el qual treballem és el complex $\mathbf{C}$, $x^2+1 = (x-i)(x+i)$, i obtenim una situació molt similar a la de l'exemple 2. El polinomi $m_f(x)$ pot ser
   - $m_f(x) = x - i$, d'on $f = iI$;
   - $m_f(x) = x + i$, d'on $f = -iI$;
   - $m_f(x) = (x-i)(x+i)$, d'on $E = E^1 \oplus E^2$. La restricció de f a E^1 és iI_{E^1}; la de f a E^2 és $-iI_{E^2}$.

   Agafant bases de E^1 i E^2 obtenim una base de E en la qual f té una matriu diagonal:

$$\left(\begin{array}{cccccc} i & & & & & \\ & \ddots & & & 0 & \\ & & i & & & \\ & & & -i & & \\ & 0 & & & \ddots & \\ & & & & & -i \end{array}\right).$$

   Si el cos sobre el qual treballem és el real $\mathbf{R}$, x^2+1 és irreductible: $m_f(x) = x^2+1$ i (4.4) no dóna cap descomposició pròpia. Podria ser, però, que aconseguíssim simplificar la matriu de f agafant vectors propis per formar una base, tal com hem estudiat al §1. Tampoc no és possible. Per què? Si k fos un valor propi, el subespai de vectors propis $\mathrm{Nuc}(f - kI) \neq \{\vec{0}\}$ seria invariant i amb polinomi mínim $x-k$. Això implicaria que $x-k$ divideix $m_f(x)$, que no és cert per a cap k. Aquest raonament que acabem de fer és general i ens dóna:

Proposició 4.5 *Si k és un valor propi de f, $(x-k) \mid m_f(x)$.*□

Exemple:

1. Sigui $f \in \mathrm{End}(E)$ tal que $f^3 = I$. El polinomi $x^3 - 1$ és un anul·lador i $m_f(x) \mid x^3 - 1$. Si el cos és $\mathbf{C}$,

$$x^3 - 1 = (x-1)(x + \frac{1+i\sqrt{3}}{2})(x + \frac{1-i\sqrt{3}}{2}).$$

Un estudi semblant al dels exemples anteriors ens dóna, per a tots els possibles $m_f(x)$, bases de E en què la matriu de f és diagonal i els valors propis són

$$\left\{1, \frac{-1+i\sqrt{3}}{2}, \frac{-1-i\sqrt{3}}{2}\right\}$$

o un subconjunt d'aquest.

Si el cos és $\mathbf{R}$, $x^3 - 1 = (x-1)(x^2+x+1)$ i, per tant, $E = E^1 \oplus E^2$. El polinomi mínim de f sobre E^1 és $x - 1$ i, per tant, f sobre E^1 és I_{E^1}. El polinomi mínim de f sobre E^2 és $x^2 + x + 1$, irreductible. Sigui $e_1, \ldots, e_r$ una base de E^1, i $e_{r+1}, \ldots, e_n$ una base de E^2. La matriu de f en la base $e_1, \ldots, e_n$ és de la forma

$$\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & A_2 \end{pmatrix}.$$

Com a l'exemple 3 es veu que cap valor propi no permet simplificar A_2.

2. Sigui $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ amb matriu

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

en la base $(1,0),(0,1)$. És fàcil veure que $m_f(x) = (x-1)^2$. El teorema (4.4) no permet descompondre E en suma de subespais invariants. Hi ha, però, un subespai invariant: el subespai de vectors propis de valor propi 1, $\langle(1,0)\rangle$. Si f tingués una matriu diagonal hauria d'ésser

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

ja que 1 és l'unic valor propi de f. Però $f \neq I$ i no pot tenir mai la matriu identitat.

3. Sigui $f \in \mathrm{End}(E)$ amb matriu

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix}$$

en la base e_1, e_2, e_3. En calcular les potències A^n es veu de seguida que $f^3 = f^2$ i, per tant, $x^3 - x^2$ és un polinomi anul·lador. Així, doncs, $m_f(x) \mid x^3 - x^2 = x^2(x-1)$.

Si $m_f(x) = x - 1$, $f = I$, que és fals.

Si $m_f(x) = x$, $m_f(f) = f = 0$, que és fals.

Si $m_f(x) = x^2$, $f^2 = 0$, que és fals.

Si $m_f(x) = (x-1)x$, $f^2 = f$, que és fals.

Així doncs, $m_f(x) = (x-1)x^2$ i $E = E^1 \oplus E^2$. Sobre E^1 f és I_{E^1}. Sobre E^2, f té polinomi mínim x^2. El càlcul de E^1 i E^2 ens dóna

$$E^1 = \operatorname{Nuc}(f - I) = \langle e_1 \rangle \qquad E^2 = \operatorname{Nuc} f^2 = \langle e_2, e_3 \rangle.$$

En la base e_1, e_2, e_3 la matriu de f és la matriu A donada. I els valors propis, quins són? Doncs són 1 i 0; els subespais de vectors propis respectius són $\langle e_1 \rangle$ i $\langle e_2 + e_3 \rangle$. Si completem aquests dos vectors fins a obtenir una base de E: $e_1, e_2 + e_3, e_3$ obtenim la matriu de f

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

que és triangular.

Tots aquests exemples estudiats ens porten de manera natural a la següent conclusió.

Teorema 4.6 (de diagonalització.) *Un endomorfisme és diagonalitzable si i només si el seu polinomi mínim descompon en factors lineals no repetits.*

DEMOSTRACIÓ: Sigui $m_f(x) = (x - a_1) \cdots (x - a_r)$, on $a_i \neq a_j$ si $i \neq j$. Per (4.4), $E = E^1 \oplus \cdots \oplus E^r$, on

$$E^i = \operatorname{Nuc}(f - a_i I)$$

és el subespai de vectors propis de valor propi a_i. Existeix, doncs, una base de vectors propis de E formada a partir de bases de cada un dels subespais $E^1, \ldots, E^r$. Recíprocament, suposem ara que E té una base formada per vectors propis:

$$e_1^1, \ldots, e_{n_1}^1, e_1^2, \ldots, e_{n_2}^2, \ldots, e_1^r, \ldots, e_{n_r}^r.$$

Suposem també que a_i és el valor propi de $e_1^i, \ldots, e_{n_i}^i$. Aleshores si posem $E^i = \langle e_1^i, \ldots, e_{n_i}^i \rangle$ es té

$$E = E^1 \oplus \cdots \oplus E^r$$

i el polinomi mínim de E^i és $(x - a_i)$. Això implica que

$$(x - a_1) \cdots (x - a_r) \mid m_f(x).$$

Ara bé, $(x-a_1)\cdots(x-a_r)$ és un anul·lador, ja que si $u = u_1+\ldots+u_r$ amb $u_i \in E^i$, $i = 1,\ldots,r$, es té

$$\begin{aligned}&(f-a_1I)\cdots(f-a_rI)(u)\\&\quad=(f-a_2I)\cdots(f-a_rI)(f-a_1I)(u_1)\\&\quad+(f-a_1I)(f-a_3I)\cdots(f-a_rI)(f-a_2I)(u_2)\\&\quad+\ldots+\\&\quad+(f-a_1I)\cdots(f-a_rI)(u_r)=\vec{0}+\vec{0}+\ldots+\vec{0}=\vec{0}.\end{aligned}$$

D'on $m_f(x) = (x-a_1)\cdots(x-a_r)$. □

VIII.5 Grau del polinomi mínim

Per la seva definició (§3) el polinomi mínim té grau $\leq n^2 = \dim \mathrm{End}(E)$. Aquesta cota és, però, molt gran a l'hora de trobar el polinomi mínim d'un endomorfisme.

Proposició 5.1 *grau* $m_f(x) \leq n$.

DEMOSTRACIÓ: Sigui $m_f(x) = m_1(x)^{n_1}\cdots m_r(x)^{n_r}$ i $E = E^1 \oplus \cdots \oplus E^r$ la descomposició de (4.4). És suficient veure que $\mathrm{gr}\, m_i(x)^{n_i} \leq \dim E^i$ per a $i = 1,\ldots,r$.

Com que $m_i(x)^{n_i}$ és el polinomi mínim de la restricció de f a E^i, hi ha un $v_i \in E^i$ tal que $m_i(f)^{n_i}(v_i) = 0$ però $m_i(f)^{n_i-1}(v_i) \neq 0$. Llavors el polinomi mínim de v_i és $m_i(x)^{n_i}$ i per (3.1) tindrem que $v_i, f(v_i), \ldots, f^{k_i-1}(v_i)$ són linealment independents, on $k_i = \mathrm{gr}\, m_i(x)^{n_i}$. Per tant, $k_i \leq \dim E^i$.□

VIII.6 El teorema de Cayley-Hamilton

A (4.5) hem vist que els zeros del polinomi característic, que designarem per $p_f(x)$, són també zeros del polinomi mínim, $m_f(x)$. Demostrarem ara el recíproc.

Proposició 6.1 *a és un zero de* $m_f(x)$ *si i només si és un zero de* $p_f(x)$.

DEMOSTRACIÓ: Si a és un zero de $m_f(x)$, $m_f(x) = (x-a)\cdot m_1(x)$. Existeix un $u \in E$ tal que $m_1(f)(u) \neq \vec{0}$ (en cas contrari el polinomi mínim seria $m_1(x)$). Aleshores el vector $w = m_1(f)(u)$ té valor propi a:

$$(f-aI)w = (f-aI)m_1(f)(u) = m_f(f)(u) = \vec{0}.$$

Per tant, a és un zero de $p_f(x)$. □

Teorema 6.2 *Si el polinomi mínim,* $m_f(x)$*, i el polinomi característic,* $p_f(x)$*, descomponen en factors lineals,* $p_f(x)$ *és un anul·lador de* f*:* $p_f(f) = 0$.

DEMOSTRACIÓ: Sigui $m_f(x) = (x-a_1)^{n_1} \cdots (x-a_r)^{n_r}$ i sigui $E = E^1 \oplus \cdots \oplus E^r$ la descomposició de (4.4). La matriu de f en una base formada per bases dels E^i és de la forma

$$A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_r \end{pmatrix}.$$

Ara bé, si $p_i(x)$ és el polinomi característic de A_i (és a dir, de la restrició de f a E^i),

$$p_f(x) = p_1(x) \cdot p_2(x) \cdots p_r(x).$$

Cada $p_i(x)$ descompon en factors lineals i els seus zeros són zeros del polinom mínim de E^i, que és $(x-a_i)^{n_i}$. Per tant, $p_i(x) = (x-a_i)^{m_i}$ amb $m_i = \dim E^i$. Per (5.1) $n_i \leq m_i$, d'on resulta que

$$m_f(x) \mid p_f(x),$$

i $p_f(x)$ és un anul·lador. □

El resultat de (6.2) val en condicions molt més generals. Per veure-ho, observem que si A és una matriu $n \times n$ sobre un cos K, podem parlar del polinomi característic de A, $p_A(x)$, i del polinomi mínim de A, $m_A(x)$, tal com ho hem fet per endomorfismes. Així $m_A(x)$ serà un polinomi de grau mínim de l'ideal $\{p(x) \in K[x] \mid p(A) = 0\}$ on, si $p(x) = a_0 + a_1 x + \ldots + a_n x^n$, posem

$$p(A) = a_0 I + a_1 A + \ldots + a_n A^n.$$

Si $p_A(x)$ i $m_A(x)$ descomponen en factors lineals, (6.2) assegura que

$$p_A(A) = 0.$$

Sigui ara A una matriu real. A serà també una matriu complexa i el polinomi característic és el mateix $p_A(x) = \det(A - xI)$. Aleshores, a $\mathbf{C}$,

$$p_A(A) = 0$$

i, naturalment, aquesta igualtat val també a $\mathbf{R}$. Aquest raonament fet per a $\mathbf{R}$ i $\mathbf{C}$, serveix per a dos cossos qualssevol $K \subset K'$, tals que tot polinomi de K' descompongui en factors lineals. Vam veure a (II.7) que si un polinomi no tenia zeros a K podíem construir un cos $K_1 \supset K$ on aquest polinomi tingués un zero. Es pot demostrar que l'aplicació successiva d'aquest procés porta a un cos $K' \supset K$ en què tot polinomi descompon en factors lineals. El cos K' es diu la *clausura algebraica* de K. Així, doncs, tot cos té una clausura algebraica i tenim en general

Teorema 6.3 (de Cayley-Hamilton) *El polinomi mínim divideix sempre el polinomi característic.*

Observació:

Aquest teorema proporciona un mètode pràctic per a calcular el polinomi mínim d'un endomorfisme f donat. Sigui A la matriu de f en una base qualsevol. Calculeu el polinomi característic $p_A(x)$, descomponeu-lo en factors irreductibles i busqueu el més petit dels seus divisors $q(x)$ tals que $q(f) = 0$ (tal com s'ha fet als exemples del §4). Aquest serà $m_f(x)$.

VIII.7 Matriu canònica (general) d'un endomorfisme

El teorema de descomposició en subespais invariants (4.4) permet reduir l'estudi d'un endomorfisme f a l'estudi de les seves restriccions a certs subespais invariants E^i. En casos molt particulars, (4.6), els espais E^i són subespais de vectors propis i podem obtenir una matriu de f diagonal. Anem ara a estudiar les restriccions de f als E^i en el cas general. Concretament, anem a descompondre cada subespai E^i en suma de subespais invariants sobre els quals l'actuació de f és molt clara: els subespais f-cíclics.

Un subespai F de E és *f-cíclic* si existeix un vector $u \in E$ tal que $F = \langle u, f(u), f^2(u), \ldots \rangle$. F és invariant per f i, per (3.1), la seva dimensió és el grau del polinomi mínim de f a u. Si aquest polinomi és

$$a_0 + a_1 x + \cdots + a_{s-1} x^{s-1} + x^s,$$

llavors $\{u, f(u), \ldots, f^{s-1}(u)\}$ és una base de F i, en aquesta base, la matriu de la restricció de f és

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & 1 & -a_{s-1} \end{pmatrix}.$$

Teorema 7.1 (segon teorema de descomposició) *Si $f \in \mathrm{End}(E)$, E és suma directa de subespais f-cíclics.*

DEMOSTRACIÓ: Per (4.4) només cal considerar el cas en què el polinomi mínim de f és una potència d'un polinomi primer: $m_f(x) = q(x)^s$, $q(x)$ primer.

Farem inducció sobre la dimensió de E. Si $\dim E = 1$, E és ja f-cíclic. Suposem el teorema cert per a tots els espais de dimensió $\leq n-1$.

Sigui $\dim E = n$. En la demostració de (5.1) vam provar que existeix un $u_0 \in E$ tal que el polinomi mínim de f a u_0 és el mateix $m_f(x)$. Sigui

$$F_0 = \langle u_0, f(u_0), f^2(u_0), \ldots \rangle$$

i $\bar{E} = E/F_0$. f indueix un endomorfisme $\bar{f}$ de $\bar{E}$,

$$\begin{array}{rccc} \bar{f} : \bar{E} & \longrightarrow & \bar{E} \\ [v] & \longmapsto & [f(v)], \end{array}$$

ben definit ja que la imatge per f de tots els representants de $[v]$ és a $[f(v)]$:

$$f([v]) = f(v + F_0) = f(v) + f(F_0) \subset f(v) + F_0 = [f(v)].$$

Ara bé, $\dim \bar{E} < n = \dim E$ i, per hipòtesi d'inducció,

$$\bar{E} = \bar{F}_1 \oplus \cdots \oplus \bar{F}_r$$

amb $\bar{F}_i$ $\bar{f}$-cíclic , $i = 1, \ldots, r$. Sigui $\bar{F}_i = \langle [u_i'], \bar{f}[u_i'], \ldots \rangle$.

Lema 7.2 *Hi ha un representant u_i de $[u_i']$ tal que si denotem per F_i el subespai $\langle u_i, f(u_i), \ldots \rangle$ la projecció*

$$\begin{array}{rccc} \pi_i : & F_i & \longrightarrow & \bar{F}_i \\ & v & \longmapsto & [v] \end{array}$$

és un isomorfisme.

Suposem, de moment, demostrat aquest lema. Aleshores

$$E = F_0 \oplus F_1 \oplus \cdots \oplus F_r.$$

Comprovem-ho: si $v \in E$, $[v] = [v_1] + \cdots + [v_r]$ amb $[v_i] \in \bar{F}_i$. El lema ens assegura aleshores que podem suposar que $v_i \in F_i$. Així doncs $E = F_0 + F_1 + \cdots + F_r$. Per veure que la suma és directa, suposem que $v \in E$ s'expressa de dues maneres com a suma de vectors de $F_0, \ldots, F_r$:
$v = v_0 + v_1 + \ldots + v_r = w_0 + w_1 + \ldots + w_r$, $v_i, w_i \in F_i$, $i = 0, 1, \ldots, r \Rightarrow$ $[v] = [v_1] + \cdots + [v_r] = [w_1] + \ldots + [w_r]$ a $\bar{E} = \bar{F}_1 \oplus \cdots \oplus \bar{F}_r \Rightarrow [v_i] = [w_i]$, $i = 1, \ldots, r$. Però, pel lema, $v_i = w_i$, $i = 1, \ldots, r$; d'on també $v_0 = w_0$.□

Només resta demostrar el lema. Fem abans dues observacions generals:

- Si $m(x)$ i $\bar{m}(x)$ són els polinomis mínims de f i $\bar{f}$ a v i a $[v]$ respectivament, $\bar{m}(x) \mid m(x)$, ja que
$$m(\bar{f})([v]) = [m(f)(v)] = [\vec{0}].$$

- El polinomi mínim de f a v divideix el polinomi mínim de f. Això implica, en el nostre cas, que aquell polinomi és una potència de $q(x)$ amb exponent $\leq s$.

DEMOSTRACIÓ: del lema. Siguin $q(x)^{s'}$ i $q(x)^{\bar{s}}$ els polinomis mínims de f i $\bar{f}$ a u_i' i $[u_i']$, $\bar{s} \leq s' \leq s$. Llavors $q(\bar{f})^{\bar{s}}([u_i']) = [\vec{0}] \Rightarrow q(f)^{\bar{s}}(u_i') \in F_0 \Rightarrow q(f)^{\bar{s}}(u_i') = a(f)(u_0)$ $\Rightarrow q(f)^{s-\bar{s}}a(f)(u_0) = q(f)^s(u_i') = \vec{0} \Rightarrow q(x)^s \mid q(x)^{s-\bar{s}}a(x)$ (ja que el polinomi mínim de u_0 és $q(x)^s$) $\Rightarrow a(x) = q(x)^{\bar{s}}b(x) \Rightarrow q(f)^{\bar{s}}(u_i') = a(f)(u_0) = q(f)^{\bar{s}}b(f)(u_0)$. El vector $u_i = u_i' - b(f)(u_0) \in u_i' + F_0 = [u_i']$ té un polinomi mínim que divideix $q(x)^{\bar{s}}$, ja que
$$q(f)^{\bar{s}}(u_i' - b(f)(u_0)) = \vec{0}.$$
Per altra banda, el polinomi mínim de u_i ha de ser múltiple del de $[u_i] = [u_i']$, que és $q(x)^{\bar{s}}$. Així doncs, el polinomi mínim de u_i és $q(x)^{\bar{s}}$, el mateix que el de $[u_i]$. Els

vectors $u_i, f(u_i), \ldots, f^{t-1}(u_i)$, on $t = \bar{s} \cdot \operatorname{grau} q(x)$, formen, per (3.1), una base de F_i. Les classes

$$[u_i], \bar{f}([u_i]), \ldots, \bar{f}^{t-1}([u_i])$$

formen, també per (3.1), una base de $\bar{F}_i$, i per tant la projecció $\pi_i : F_i \longrightarrow \bar{F}_i$ és un isomorfisme. □

Fins a quin punt és única la descomposició obtinguda a (7.1)? Suposem que

$$E = G_0 \oplus G_1 \oplus \cdots \oplus G_m,$$

on els subespais G_i són f-cíclics i el polinomi mínim de la restricció de f a G_i és $q_i(x)^{n_i}$ amb $q_i(x)$ irreductible. Agrupem els sumands que corresponen a potències del mateix $q_i(x)$. Per exemple, suposem $q_0(x) = q_1(x) = \ldots = q_s(x)$ i considerem $E^0 = G_0 \oplus \cdots \oplus G_s$. El polinomi mínim de la restricció de f a E^0 és $q_0(x)^t$, on $t = \max(n_0, \ldots, n_s)$. Agrupant d'aquesta forma els G_i obtenim una descomposició de E,

$$E = E^0 \oplus E^1 \oplus \cdots \oplus E^r,$$

que és precisament la de (4.4). Així, doncs, els subespais E^i estan unívocament determinats i les diferents descomposicions de E en subespais f-cíclics correspondran a les diferents descomposicions dels E^i.

Considerem, doncs, el cas en què el polinomi mínim de $f \in \operatorname{End}(E)$ és $q(x)^s$, $q(x)$ primer. La primera observació que cal fer és que no podem aspirar a demostrar la unicitat de la descomposició de (7.1). Ho veurem amb un exemple.

Exemple:

Si $f = kI_E$, qualsevol base $u_1, \ldots, u_n$ dóna lloc a una descomposició en subespais f-cíclics:

$$E = \langle u_1 \rangle \oplus \cdots \oplus \langle u_n \rangle.$$

Anem a veure, però, que, en totes les possibles descomposicions de E en subespais f-cíclics el nombre n_t de subespais als quals correspon un cert polinomi mínim $q(x)^t$ és el mateix. Recordem que estem considerant el cas en què el polinomi mínim de f és $q(x)^s$. Suposem que

$$E = F_0 \oplus \cdots \oplus F_r$$

és una descomposició de E en suma de subespais f-cíclics F_i. Sigui $q(x)^{s_i}$ el polinomi mínim de la restricció de f a F_i ($s_i \leq s$).

Tenim:

a) $q(f)^t(F_i) \subset F_i$; d'on

$$q(f)^t(E) = q(f)^t(F_0) \oplus \cdots \oplus q(f)^t(F_r),$$

i, per tant,

$$\dim\left(q(f)^t(E)\right) = \sum_{i=0}^{r} \dim(q(f)^t(F_i)).$$

b) Si $s_i > t$, el polinomi mínim de la restricció de f a $q(f)^t(F_i)$ és $q(x)^{s_i - t}$. Llavors,

$$\dim(q(f)^t(F_i)) = (s_i - t) \cdot \operatorname{gr} q(x) = \dim F_i - t \cdot \operatorname{gr} q(x).$$

c) Si $s_i \leq t$, $q(f)^t(F_i) = \{\vec{0}\}$. Aleshores,

$$\dim(q(f)^t(F_i)) = 0 = \dim F_i - s_i \cdot \operatorname{gr} q(x).$$

Substituint les expressions de b) i c) en el sumatori de a) obtenim

$$\dim(q(f)^t(E)) = \dim E - \sum_{i=0}^{r} \text{mín}(s_i, t) \cdot \operatorname{gr} q(x).$$

Observem ara que

$$\text{mín}(s_i, t) - \text{mín}(s_i, t-1)) = \begin{cases} 0 & \text{si } s_i \leq t-1 \\ 1 & \text{si } s_i \geq t, \end{cases}$$

d'on $\sum_{i=0}^{r}(\text{mín}(s_i, t) - \text{mín}(s_i, t-1)) = n_t + n_{t+1} + \cdots + n_s$. Denotem per q_t la dimensió de $q(f)^t(E)$. Tenim llavors

$$q_{t-1} - q_t = (n_t + n_{t+1} + \cdots + n_s) \cdot \operatorname{gr} q(x),$$

d'on resulta que

$$n_t = \frac{1}{\operatorname{gr} q(x)}(q_{t-1} - 2q_t + q_{t+1}).$$

Aquesta expressió de n_t no depèn de la descomposició de E considerada.

Observació:

Es compleix

$$\{\vec{0}\} \subset \operatorname{Nuc} q(f) \subset \operatorname{Nuc} q(f)^2 \subset \ldots \subset \operatorname{Nuc} q(f)^s = \operatorname{Nuc} q(f)^{s+1} = \ldots = E.$$

Sigui $\bar{q}_t = \dim \operatorname{Nuc} q(f)^t = n - q_t$ i designem per

$$p_t = \dim\left(\operatorname{Nuc} q(f)^t / \operatorname{Nuc} q(f)^{t-1}\right) = \bar{q}_t - \bar{q}_{t-1}.$$

Resulta, aleshores, que

$$n_t = \frac{1}{\operatorname{gr} q(x)}(p_t - p_{t+1}).$$

VIII.8 Matriu canònica de Jordan

Suposem ara que el polinomi mínim de f descompon en factors lineals. Sigui

$$E = F_0 \oplus \cdots \oplus F_r$$

la descomposició de (7.1) en subespais f-cíclics i $(x-a_i)^{s_i}$ el polinomi mínim de la restricció de f a F_i. Triem per a cada F_i un vector u_i amb polinomi mínim $(x-a_i)^{s_i}$. Aleshores,

$$u_i, (f-a_iI)(u_i), \ldots, (f-a_iI)^{s_i-1}(u_i)$$

són linealment independents. Ara bé, per (3.1), $\dim F_i = s_i$ i per tant aquests vectors formen una base. La matriu de la restricció de f a F_i en aquesta base és

$$J(a_i, s_i) = \left.\begin{pmatrix} a_i & 0 & \cdots & 0 & 0 \\ 1 & a_i & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_i & 0 \\ 0 & 0 & \cdots & 1 & a_i \end{pmatrix}\right\} s_i.$$

Tenim així com a conseqüència de (7.1)

Teorema 8.1 *Si el polinomi mínim de $f \in \mathrm{End}(E)$ descompon en factors lineals, existeix una base de E en la qual la matriu de f és de la forma*

$$J = \begin{pmatrix} \boxed{J(a_0, s_0)} & & & 0 \\ & \boxed{J(a_1, s_1)} & & \\ & & \ddots & \\ 0 & & & \boxed{J(a_r, s_r)} \end{pmatrix}. \square$$

La matriu J es diu la *matriu canònica de Jordan de f.*

Observem que els elements a_i que apareixen a la diagonal de la matriu canònica de Jordan són els valors propis de f (possiblement repetits). Per obtenir una base de E en què la matriu de f sigui la matriu canònica de Jordan procedirem de la següent forma.

Siguin

$$p(x) = (x-\lambda_1)^{\alpha_1} \cdots (x-\lambda_r)^{\alpha_r}$$

i

$$m(x) = (x-\lambda_1)^{s_1} \cdots (x-\lambda_r)^{s_r}$$

els polinomis característic i mínim de f. Recordem que $\alpha_1 + \cdots + \alpha_r = n = \dim E$ i $s_i \leq \alpha_i$ $\forall i$. Els enters s_i estan caracteritzats pel fet que

$$\{\vec{0}\} \subset \operatorname{Nuc}(f - \lambda_i I) \subset \operatorname{Nuc}(f - \lambda_i I)^2 \subset$$
$$\subset \ldots \subset \operatorname{Nuc}(f - \lambda_i I)^{s_i - 1} \subsetneq \operatorname{Nuc}(f - \lambda_i I)^{s_i}$$

i $\forall t \geq s_i$

$$\operatorname{Nuc}(f - \lambda_i)^{s_i} = \operatorname{Nuc}(f - \lambda_i)^t = E^i.$$

A més a més,

$$E = E^1 \oplus \cdots \oplus E^r.$$

La base que busquem és unió de bases convenients dels subespais invariants E_i. Restringint-nos a aquests espais, podem suposar que el polinomi característic de $f \in \operatorname{End}(E)$ és $(x - \lambda)^n$ i el polinomi mínim $(x - \lambda)^s$, $s \leq n$. Per a qualsevol $u \in E$ designarem $(f - \lambda I)(u)$ per $q(u)$. Considerem el requadre de vectors de la pàgina 160. $u_{11}, \ldots, u_{1k_1}$ determinen classes que formen una base de $\operatorname{Nuc}(f - \lambda I)^s / \operatorname{Nuc}(f - \lambda I)^{s-1}$. És fàcil veure que $u_{11}, \ldots, u_{1k_1}$ són linealment independents i que $q(u_{11}), \ldots, q(u_{1k_1}) \in \operatorname{Nuc}(f - \lambda I)^{s-1}$ determinen classes linealment independents a $\operatorname{Nuc}(f - \lambda I)^{s-1} / \operatorname{Nuc}(f - \lambda I)^{s-2}$. Els vectors $u_{21}, \ldots, u_{2k_2}$ són vectors representants de classes que, juntament amb les anteriors, formen una base de $\operatorname{Nuc}(f - \lambda I)^{s-1} / \operatorname{Nuc}(f - \lambda I)^{s-2}$. Repetim ara aquest procés fins obtenir una base de $\operatorname{Nuc}(f - \lambda I)$ formada pels vectors situats a l'última fila del requadre.

Tenim aleshores:

i) El conjunt de tots els vectors que apareixen en el requadre formen una base de E.

ii) El nombre de columnes és la multiplicitat del valor propi λ.

iii) El nombre de files és l'exponent del polinomi mínim.

iv) El nombre de vectors a cada fila és

$$\dim(\operatorname{Nuc}(f - \lambda I)^t / \operatorname{Nuc}(f - \lambda I)^{t-1}) = \bar{q}_t - \bar{q}_{t-1} = p_t.$$

v) El nombre de matrius $J(\lambda, t)$ que apareixen en la matriu canònica de Jordan de f és $n_t = p_t - p_{t+1}$, que és la diferència de vectors en files consecutives.

vi) Cada columna és la base d'un subespai f-cíclic de la descomposició de E.

Observació:

$n_s \geq 1$ sempre, ja que $p_{s+1} \geq 1$. Això vol dir que sempre apareix almenys una matriu $J(\lambda, s)$ de la màxima dimensió.

$E = \mathrm{Nuc}(f-\lambda I)^s$	u_{11}	$\cdots$	u_{1k_1}
$\mathrm{Nuc}(f-\lambda I)^{s-1}$	$q(u_{11})$	$\cdots$	$q(u_{1k_1})$	u_{21}	$\cdots$	u_{2k_2}
$\mathrm{Nuc}(f-\lambda I)^2$	$q^{s-2}(u_{11})$	$\cdots$	$q^{s-2}(u_{1k_1})$	$q^{s-3}(u_{21})$	$\cdots$	$q^{s-3}(u_{2k_2})$	$\cdots$	$u_{s-1,1}$	$\cdots$	$u_{s-1,k_{s-1}}$
$\mathrm{Nuc}(f-\lambda I)$	$q^{s-1}(u_{11})$	$\cdots$	$q^{s-1}(u_{1k_1})$	$q^{s-2}(u_{21})$	$\cdots$	$q^{s-2}(u_{2k_2})$	$\cdots$	$q(u_{s-1,1})$	$\cdots$	$q(u_{s-1,k_{s-1}})$	u_{s1}	$\cdots$	u_{sk_s}

Exemple:

Considerem un endomorfisme $f \in \mathrm{End}(E)$ amb matriu

$$\begin{pmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & 3 & 1 & 1 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}.$$

El seu polinomi característic és $(2-x)^5$. El rang de $(f-2I)$ és 2 i, per tant, $\dim \mathrm{Nuc}(f-2I) = 3$. A més a més,

$$(f-2I)^2 = 0,$$

i el polinomi minim de f és $(x-2)^2$. El requadre considerat anteriorment té, en aquest cas, dues files i tres elements a la fila inferior:

u_{11}	u_{12}	
$q(u_{11})$	$q(u_{12})$	u_2

Podem agafar $u_{11} = (1,0,0,0,0)$, $u_{12} = (0,0,0,1,0)$. Llavors,

$$\begin{aligned} q(u_{11}) &= (f-2I)(u_{11}) = (-1,1,0,0,0) \\ q(u_{12}) &= (f-2I)(u_{12}) = (-1,1,0,-1,1) \end{aligned}$$

són dos vectors propis que, juntament amb $u_2 = (0,0,1,0,0)$, formen una base de vectors propis. En la base

$$u_{11},\ q(u_{11}),\ u_{12},\ q(u_{12}),\ u_2,$$

la matriu de f és

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

El problema de trobar la matriu canònica de Jordan d'un endomorfisme f donat queda, doncs, resolt, si el polinomi mínim de f descompon en factors lineals:

$$m(x) = (x-\lambda_1)^{s_1} \cdots (x-\lambda_r)^{s_r}.$$

Només ens cal conèixer per a cada valor propi λ_i els nombres n_t corresponents ($1 \leq t \leq s_i$), els quals ens indiquen quants subespais f-cíclics de dimensió t apareixen (submatrius amb el valor propi λ_i a la diagonal i uns a sota). En particular, si $s_i = 1$, el subespai invariant corresponent és un subespai de vectors propis. L'endomorfisme és diagonalitzable si i només si $s_1 = \ldots = s_r = 1$ (4.6).

VIII.9 Nota històrica

El 1858 Arthur Cayley (1821–1895) va enunciar en general el teorema que ara es coneix amb el nom de Cayley-Hamilton, demostrant-lo per matrius 3×3 i introduint el polinomi característic d'una matriu i les seves arrels (valors propis).

Henry Taber (1860–?) va enunciar la proposició 1.2, i, en particular, que el terme independent del polinomi característic és el determinant de la matriu, introduint també la traça. La demostració del teorema la va fer William Henry Metzler (1863–?) el 1891.

Georg Ferdinand Frobenius (1849–1917) va introduir el 1878 el polinomi mínim i Kurt Hensel (1861–1941) en va demostrar, el 1904, les principals propietats; en particular, que qualsevol altre anul·lador és múltiple del polinomi mínim.

Utilitzant el concepte de matrius equivalents i el polinomi característic, Camille Jordan (1838–1922) va demostrar el 1870 que tota matriu és equivalent a una de forma canònica (la forma canònica de Jordan).

El desenvolupament algebraic dels temes d'aquest capítol fou de gran transcendència per a la física. Com va profetitzar Peter G. Tait (1831–1901), "Cayley is forging the weapons for future generations of physicists".

VIII.10 Exercicis

1. Sigui A la matriu de $M_{n\times n}(K)$ formada íntegrament per uns. Calculeu els polinomis característic i mínim de A. Proveu que A és diagonalitzable i trobeu una matriu diagonal D i una invertible M tals que $A = MDM^{-1}$.

2. a) Determineu totes les matrius $A \in M_{2\times 2}(\mathbf{R})$ tals que

$$\text{i) } A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \qquad \text{ii) } A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix};$$
$$\text{iii) } A^2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; \quad \text{iv) } A^2 = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}.$$

b) Determineu totes les matrius $A \in M_{4\times 4}(\mathbf{R})$ tals que

$$A^2 - 3A + 2I = 0.$$

c) Determineu totes les matrius $A \in M_{n\times n}(\mathbf{R})$ tals que $A^2 = A$.

3. Donada

$$A = \begin{pmatrix} 5/2 & -1 \\ 3 & -1 \end{pmatrix},$$

calculeu A^{1438} i $\lim_{n\to\infty} A^n$.

4. Sigui $e_1, \ldots, e_n$ una base de l'espai vectorial E i $f \in \mathrm{End}(E)$ tal que

$$f(e_1) = \ldots = f(e_n) = \sum_{i=1}^{n} a^i e_i.$$

Demostreu que f és diagonalitzable si i només si $\sum_{i=1}^{n} a^i \neq 0$.

5. Determineu la forma general de les matrius que commuten amb les matrius diagonals i de les que commuten amb les diagonalitzables.

6. Demostreu que $f \in \mathrm{End}_{\mathbf{C}}(E)$ és diagonalitzable si i només si tot subespai invariant per f admet un complementari també invariant per f.

7. Construïu un endomorfisme f de $\mathbf{C}^3$ tal que

$$p_f(x) = m_f(x) = x^2(x-a), a \neq 0.$$

Demostreu que si u és un vector propi de valor propi a i $v \in \mathrm{Nuc}\, f^2 - \mathrm{Nuc}\, f$, aleshores $u, v, f(v)$ és una base de $\mathbf{C}^3$. Trobeu la matriu de f en aquesta base i calculeu f^n.

8. Sigui $f \in \mathrm{End}_{\mathbf{R}}(E)$. Demostreu que si $\dim E$ és senar, aleshores f té algun valor propi i que si f és parell i $\det f < 0$, f té almenys dos valors propis. Doneu un exemple d'un endomorfisme sense valors propis.

9. Demostreu que si $f \in \mathrm{End}(E)$ és diagonalitzable i F és un subespai invariant per f, aleshores $f \mid F$ també és diagonalitzable.

10. Demostreu que si $f, g \in \mathrm{End}(E)$ commuten, els subespais de vectors propis de g són invariants per f i recíprocament.

11. Es diu que dos endomorfismes $f, g \in \mathrm{End}(E)$ són simultàniament diagonalitzables si podem trobar una base de E respecte a la qual les dues matrius de f i g siguin diagonals.

a) Demostreu que f i g són simultàniament diagonalitzables si i només si són diagonalitzables i commuten.
(Indicació: utilitzeu els dos exercicis anteriors).

b) Diagonalitzeu simultàniament els dos endomorfismes de $\mathbf{R}^3$:

$$\begin{array}{ll} & f(x,y,z) = (x+y+z, 2x+5y+2z, -2x-5y-2z) \\ \text{i} & g(x,y,z) = (-2y-2z, 0, 2y+2z). \end{array}$$

12. Donada $A \in M_{n\times n}(\mathbf{R})$, considerem l'equació diferencial

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix};$$

és a dir, busquem n funcions diferenciables $x_i : \mathbf{R} \to \mathbf{R}$ que satisfan una determinada relació entre elles i les seves derivades.

a) Convenceu-vos que si A és diagonal (i, per tant, si A és diagonalitzable) sabem resoldre qualsevol equació diferencial d'aquest tipus.

b) Resoleu les següents equacions diferencials:

$$\text{i)} \begin{cases} x' = y \\ y' = x \end{cases} ; \qquad \text{ii)} \begin{cases} x' = -x - 2z \\ y' = 6x + y + 6z \\ z' = x + 2z \end{cases} ; \qquad \text{iii)} \ y'' - y' = y \ .$$

(Indicació: a iii) introduïu la nova variable $z = y'$).

13. Siguin $(a_n), (b_n), \ldots,$ m successions tals que

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \\ \vdots \end{pmatrix} = A \begin{pmatrix} a_n \\ b_n \\ \vdots \end{pmatrix}$$

amb $A \in M_{m \times m}(K)$. Trobeu una expressió (no recurrent) del terme general de cada una d'aquestes successions quan A és diagonal. Com es pot trobar aquest terme general quan A és diagonalitzable?

Apliqueu-ho als següents casos:

a) Trobeu totes les successions (a_n) tals que $a_{n+1} = 2a_n + a_{n-1}$. (Indicació: introduïu una nova successió $b_n = a_{n-1}$).

b) Trobeu el terme general de la successió de Fibonacci 0, 1, 1, 2, 3, 5, 8, 13, ...

c) Discutiu la convergència de les successions complexes (a_n), (b_n) donades per

$$a_{n+1} = \alpha(a_n + \sqrt{3}b_n) \quad \text{i} \quad b_{n+1} = \alpha(-\sqrt{3}a_n + b_n)$$

segons els valors de α, a_0 i b_0.

14. Quines de les matrius

$$A = \begin{pmatrix} 1 & -1 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \ C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

poden ésser associades al mateix endomorfisme? Per a aquestes busqueu la matriu del canvi de base.

15. Sigui $f \in \mathrm{End}(E)$, $\dim E = n$. Demostreu:

a) Existeix un $k \in \mathbf{N}$ tal que $\mathrm{Nuc}\, f^k = \mathrm{Nuc}\, f^{k+i}$, $\forall i \geq 0$.

b) $E = \mathrm{Nuc}\, f^k \otimes \mathrm{Im}\, f^k$ i $f \mid \mathrm{Im}\, f^k$ és un automorfisme de $\mathrm{Im}\, f^k$.

16. Estudieu els endomorfismes f que cumpleixen $f^3 = a^2 f$, $a \in \mathbf{R}$.

17. Un endomorfisme $f \in \mathrm{End}(E)$ es diu *nilpotent* si existeix un $n \in \mathbf{N}$ tal que $f^n = 0$. Si f és nilpotent, demostreu:

a) traça $f = 0$;

b) $\det(f + I) = 1$;

c) $\forall g \in \mathrm{End}(E)$ que commuti amb f, $\det(f + g) = \det g$ (Indicació: considereu per separat els casos g invertible o no);

d) Si $g \in \mathrm{End}(E)$, aleshores $E = F \oplus G$ amb $g \mid G$ isomorfisme i $g \mid F$ nilpotent (descomposició de Fitting).

18. Sigui $f : \mathbf{R}_2[x] \longrightarrow \mathbf{R}_2[x]$ l'endomorfisme que fa correspondre $f(p) = p + p'$ a cada polinomi real p de grau més petit que 3.

a) Trobeu la forma canònica de Jordan de f.

b) Demostreu que f^{-1} és una expressió polinòmica en f.

c) Trobeu la matriu de f^{-1} en la base $1, x, x^2$. (Indicació: feu servir b)).

19. Sigui $q(x)$ un polinomi qualsevol. Demostreu que si el polinomi característic d'un endomorfisme f és $(x - a_1)^{n_1} \cdots (x - a_r)^{n_r}$, el polinomi característic de $q(f)$ és $(x - q(a_1))^{n_1} \cdots (x - q(a_r))^{n_r}$.

20. Trobeu la forma canònica de Jordan de l'endomorfisme f de $M_{2\times 2}(\mathbf{C})$ que envia $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ a $\begin{pmatrix} 2b + 5c - 6d & -a + 3b + 4c - 5d \\ -4c + 9d & -4c + 4d \end{pmatrix}$.

VIII.11 Exercicis de programar

21. Prepareu un programa que calculi els coeficients del polinomi característic d'una matriu $A \in M_{n\times n}(\mathbf{R})$ $(n \leq 5)$ fent servir l'expressió donada a la proposició 1.2.

22. Prepareu un programa que donada una matriu $A \in M_{n\times n}(\mathbf{R})$ calculi el polinomi mínim pel següent procediment:

Calculeu

$$r_k = \dim\langle I, A, A^2, \ldots, A^k\rangle \leq \dim M_{n\times n}(\mathbf{R}), \quad k = 1, 2, 3, \ldots.$$

Quan trobeu el primer $k \leq n$ tal que $r_k = k$, resoleu el sistema d'equacions

$$A^k = x_0 I + x_1 A + \ldots + x_{k-1} A^{k-1}$$

triant com a matriu del sistema aquella que havia donat $r_{k-1} = k$. S'obté una solució única i els valors de les x_i són els coeficients del polinomi mínim.

23. Donada $A \in M_{2\times 2}(\mathbf{R})$, calculeu els seus valors propis k_1,k_2 (que poden ésser reals o complexos).

- Si $k_1 \neq k_2$, trobeu una base de vectors propis.
- Si $k_1 = k_2$ i A no és una homotècia, no pot ésser diagonalitzable. Trobeu l'únic subespai de vectors propis.

24. Donada $A \in M_{3\times 3}(\mathbf{R})$, calculeu tots els seus valors propis pel següent procediment:

El polinomi característic ha de tenir almenys una arrel real α. Trobeu-la, dividiu per $x - \alpha$ i calculeu les dues arrels restants.

Per a un càlcul aproximat de α podeu fer servir el següent mètode: si $p(x) = a_0 + a_1x + a_2x^2 - x^3$ és el polinomi característic de A, totes les seves arrels són a l'interval $(-b, b)$ on $b = 1 + |a_0| + |a_1| + |a_2|$. Per tant, $p(-b) \cdot p(b) < 0$. Anem dividint successivament l'interval per la meitat quedant-nos a la part que contingui un canvi de signe. (Els càlculs són més senzills si es posa $p(x) = a_0 + x(a_1 + x(a_2 - x))$).

25. Donada $A \in M_{n\times n}(\mathbf{R})$, de la qual coneixem un valor propi real k, trobeu una base del subespai de vectors propis corresponents. (Indicació: plantegeu el sistema homogeni $(A - kI)x = \vec{0}$ i resoleu-lo amb el programa de l'exercici VII.12).

A la pràctica, per a n gran , els valors propis reals d'una matriu $A \in M_{n\times n}(\mathbf{R})$ no es busquen a partir del polinomi característic, sinó que s'aproximen per mètodes iteratius. Descriurem a continuació un d'ells i la resta al capítol XII.

26. Mètode LU

Escrivim $A = LU$ com a l'exercici IV.17. Sigui $A_1 = UL$. L'algorisme consisteix a repetir aquest càlcul: descompondre $A_k = L_kU_k$ i fer $A_{k+1} = U_kL_k$, $k = 1, 2, \ldots$

Les matrius que es van obtenint són totes equivalents a A, ja que $A_{k+1} = (L_k)^{-1}A_kL_k \ \forall k$.

Sota certes hipòtesis, la successió $\{A_k\}$ tendeix a una matriu triangular superior (i, per tant, a la diagonal quedaran els valors propis de A).

Aquestes hipòtesis són tècniques i no hi ha manera de comprovar-les *a priori*. L'algorisme fallarà, doncs, en alguns casos. D'entrada hi ha dues condicions necessàries òbvies:

a) La descomposició LU ha d'ésser possible a cada pas.

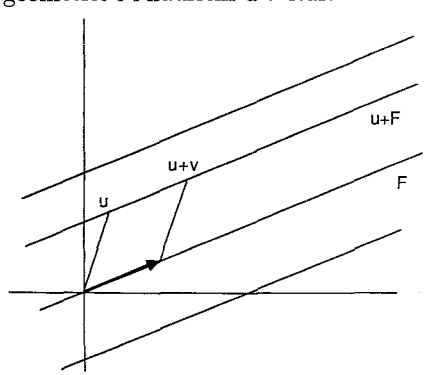
b) Tots els valors propis de A han d'ésser reals (p.e. si la matriu A és simètrica; veure XI.7.3).

Capítol IX

Espais afins

Tots nosaltres estem ja acostumats, a hores d'ara, a associar la "recta", el "pla" i l'"espai" amb $\mathbf{R}$, $\mathbf{R}^2$ i $\mathbf{R}^3$, respectivament, i tenim, fins i tot, una certa visió geomètrica de $\mathbf{R}^n$ (IV.1. Exemple 2). Un subespai F de dimensió 1 és, "geomètricament", una recta que passa per $(0,0)$; les rectes paral·leles a aquesta són, precisament, els subconjunts $u+F$ de $\mathbf{R}^2$ (recordem que la suma de vectors correspon geomètricament a la "llei del paral·lelogram"). De la mateixa manera, les rectes (plans) de $\mathbf{R}^3$ són els subconjunts $u+F$, on F és un subespai vectorial de dimensió 1 (2). Podem considerar, per tant, la "geometria afí" com l'estudi de $\mathbf{R}^n$ i dels seus subconjunts $u+F$.

El fet d'identificar l'estudi de la geometria amb l'estudi de l'espai vectorial $\mathbf{R}^n$ té, però, inconvenients. Al pla, per exemple, no hi ha cap criteri que permeti diferenciar una recta o una família de rectes entre totes; totes les rectes tenen exactament les mateixes característiques i una elecció seria, per tant, totalment arbitrària. A $\mathbf{R}^2$, en canvi, els subconjunts $u+F$ són de dos tipus ben diferenciats: uns són subespais vectorials ($\vec{0}+F=F$) i els altres no. Aquesta diferenciació no correspon a cap fet geomètric i l'hauríem d'evitar.



Si no hi ha cap motiu per a escollir una entre totes les rectes que passen per un punt-origen, fem que qualsevol punt pugui ésser aquest origen. Així, un *pla afí* serà, per definició, un conjunt A i, per a cada $p \in A$, una aplicació bijectiva

$$\varphi_p : A \longrightarrow \mathbf{R}^2$$

tal que $\varphi_p(p) = (0,0)$. Naturalment les aplicacions φ_p no poden ésser totalment arbitràries; haurem d'exigir algunes condicions que les relacionin. (De fet imposarem només una condició).

La geometria a la qual ens hem referit fins aquí és la "geometria afí real". En principi, no hi ha cap inconvenient en el fet de considerar, en lloc de $\mathbf{R}$, qualsevol cos K i, en lloc de $\mathbf{R}^2$ (o $\mathbf{R}^n$), qualsevol espai vectorial sobre K.

IX.1 Definició d'espai afí

Un *espai afí sobre un cos* K és un conjunt $A \neq \emptyset$, un espai vectorial E i una aplicació

$$\varphi : A \times A \longrightarrow E$$

que compleix:

1.
$$\begin{array}{rccc} \varphi_p : & A & \longrightarrow & E \\ & q & \longmapsto & \varphi(p,q) \end{array} \quad \text{és bijectiva } \forall p \in A.$$

2.
$$\varphi(p,q) + \varphi(q,r) = \varphi(p,r) \qquad \forall p,q,r \in A.$$

Escriurem

$$\varphi(p,q) = \overrightarrow{pq}$$

i direm que p i q són, respectivament, *l'origen* i *l'extrem* del vector $\overrightarrow{pq}$. Amb aquesta notació, la condició 2 diu que

$$\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}.$$

Els elements de A es diuen *punts*. E es diu *l'espai vectorial associat a* A, i definim la *dimensió de* A com la dimensió de E.

Exemple:

Sigui K un cos i $A = K^n$, $E = K^n$,

$$\begin{array}{rcl} \varphi : A \times A & \longrightarrow & E \\ ((a_1,\ldots,a_n),(b_1,\ldots,b_n)) & \longmapsto & (b_1 - a_1,\ldots,b_n - a_n). \end{array}$$

(A, E, φ) és un espai afí que denominem *espai afí estàndard de dimensió* n *sobre* K.

Proposició 1.1 *a)* $\overrightarrow{pq} = \vec{0} \Leftrightarrow p = q$.

b) $\overrightarrow{pq} = -\overrightarrow{qp} \qquad \forall p, q \in A$.

c) $\overrightarrow{pq} = \overrightarrow{rs} \Leftrightarrow \overrightarrow{pr} = \overrightarrow{qs}$.

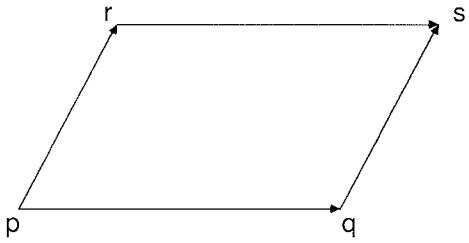
DEMOSTRACIÓ: La condició 2. a la definició d'espai afí ens diu que $\overrightarrow{pp} + \overrightarrow{pp} = \overrightarrow{pp}$. D'aquí resulta $\overrightarrow{pp} = \vec{0}$. Recíprocament, $\overrightarrow{pq} = \vec{0}$ implica $\varphi_p(q) = \vec{0}$; però com que $\varphi_p(p) = \overrightarrow{pp} = \vec{0}$, de la bijectivitat de φ_p resulta $p = q$, que demostra a).

Per provar b) observem que

$$\overrightarrow{pq} + \overrightarrow{qp} = \overrightarrow{pp} = \vec{0} \quad ; \text{és a dir, } \overrightarrow{pq} = -\overrightarrow{qp}.$$

Finalment, per demostrar c) posem

$$\overrightarrow{pr} = \overrightarrow{pq} + \overrightarrow{qr} \ , \ \overrightarrow{qs} = \overrightarrow{qr} + \overrightarrow{rs}.$$



D'aquí resulta que $\overrightarrow{pr} = \overrightarrow{qs}$ si i només si $\overrightarrow{pq} = \overrightarrow{rs}$.□

IX.2 Translacions. Una altra definició d'espai afí

Sigui (A, E, φ) un espai afí. Donat $u \in E$, anomenarem *translació de vector* u l'aplicació:

$$\begin{array}{rccc} T_u: & A & \longrightarrow & A \\ & p & \longmapsto & \varphi_p^{-1}(u); \end{array}$$

és a dir, $T_u(p)$ és un punt q tal que $\overrightarrow{pq} = u$.

Exemple:

$T_{\vec{0}}$ = Identitat.

Proposició 2.1 *a)* $\forall u \in E$, T_u *és bijectiva.*

b) Si $\exists p \in A$ *tal que* $T_u(p) = T_v(p)$, *llavors* $u = v$.

c) $T_u \circ T_v = T_{u+v}$

d) $T_{-u} = T_u^{-1}$.

e) Donats $p, q \in A$, *existeix un* $u \in E$ *i només un tal que* $T_u(p) = q$.

DEMOSTRACIÓ: T_u és injectiva, ja que

$$T_u(p) = T_u(p') = q \quad \Rightarrow \quad \overrightarrow{pq} = u = \overrightarrow{p'q} \Rightarrow$$
$$\Rightarrow \quad \varphi_q(p) = \overrightarrow{qp} = \overrightarrow{qp'} = \varphi_q(p') \Rightarrow p = p'.$$

T_u és exhaustiva, ja que donat $q \in A$, el punt $p = \varphi_q^{-1}(-u)$ compleix $\overrightarrow{qp} = -u$; per tant, $\overrightarrow{pq} = u$ i $q = T_u(p)$.

Per provar b) observem que de $T_u(p) = T_v(p) = q$ es dedueix $\overrightarrow{pq} = u$ i $\overrightarrow{pq} = v$, i per tant $u = v$.

Sigui $p \in A$ i posem $q = T_v(p)$, $r = T_u(q)$. Aleshores $T_u \circ T_v(p) = r$. Ara bé, $\overrightarrow{pr} = \overrightarrow{pq} + \overrightarrow{qr} = u + v$ i, per tant, $T_u \circ T_v(p) = T_{u+v}(p)$. Aquest raonament es pot fer per a tot $p \in A$ i ens demostra c).

d) resulta de c) i de $T_{\vec{0}} = I$.

A e), $T_u(p) = q \Leftrightarrow u = \overrightarrow{pq}$ i, per tant, aquest és el vector demanat.□

L'últim apartat de la proposició (2.1) ens diu que les translacions d'un espai afí A, $\{T_u; u \in E\}$ determinen l'aplicació $\varphi : A \times A \longrightarrow E$; $\varphi(p, q)$ és precisament el vector l'existència del qual està assegurada per e). Però encara es compleix més, com ho prova la proposició següent.

Proposició 2.2 *Donat un conjunt* A, *un espai vectorial* E *i una família d'aplicacions* $\mathcal{T} = \{T_u : A \longrightarrow A; \forall u \in E\}$ *que compleixen*

1. $T_u \circ T_v = T_{u+v}$;
2. $\forall p, q \in A$, *existeix un* $u \in E$ *i només un tal que* $T_u(p) = q$;

aleshores existeix un únic espai afí (A, E, φ) *tal que* $\mathcal{T}$ *és precisament el seu conjunt de translacions.*

DEMOSTRACIÓ: Naturalment φ ha d'ésser l'aplicació

$$\varphi : A \times A \longrightarrow E$$

tal que $\varphi(p, q) = u$ és el vector donat per 2. Només hem de comprovar, per tant, que (A, E, φ) és un espai afí.

φ_p és injectiva, ja que

$$\varphi_p(q) = \varphi_p(q') = u \Rightarrow q = T_u(p) = q'.$$

φ_p és exhaustiva, ja que l'antiimatge de $u \in E$ és $q = T_u(p)$.

Tenim també

$$\varphi(p,q)=u,\ \varphi(q,r)=v \ \Rightarrow\ T_u(p)=q,\ T_v(q)=r \Rightarrow r=T_v\circ T_u(p)=T_{v+u}(p)\Rightarrow$$
$$\Rightarrow\ \varphi(p,r)=u+v=\varphi(p,q)+\varphi(q,r).$$

Comprovem finalment que $\mathcal{T}$ és el conjunt de translacions de (A,E,φ) : si $T'_u : A \longrightarrow A$ és l'aplicació $T'_u(p)=\varphi_p^{-1}(u)$, llavors

$$T'_u(p)=q \Leftrightarrow \varphi(p,q)=u \Leftrightarrow T_u(p)=q.$$

Per tant, $T'_u=T_u\ \forall u\in E$. □

Aquesta proposició ens diu que un *espai afí* també es pot definir com un conjunt A, un espai vectorial E i una família d'aplicacions $\mathcal{T}=\{T_u : A\longrightarrow A;\ \forall u\in E\}$ que compleixin les dues propietats de (2.2).

IX.3 Varietats lineals

Sigui (A,E,φ) un espai afí. Sigui $a\in A$ i F un subespai vectorial de E. Es diu *varietat lineal* que *passa per* a i té la *direcció* F el subconjunt de A

$$\{b\in A;\ \overrightarrow{ab}\in F\}.$$

Per indicar $u=\overrightarrow{ab}$, farem servir les notacions

$$u=b-a \qquad b=a+u.$$

Amb aquesta notació, per exemple, $T_u(a)=a+u$.

Una varietat lineal és, doncs, un conjunt del tipus $\{b\in A;\ b=a+u,\ u\in F\}$, que designarem per

$$a+F.$$

Observem que $A=a+E$ és també una varietat lineal.

La *dimensió* d'una varietat lineal $a+F$ és la dimensió de la seva direcció F. Les varietats de dimensió 0 són els punts de A. Les varietats de dimensions 1, 2 i $(n-1)$ es diuen *rectes, plans* i *hiperplans*, respectivament $(n=\dim E)$.

Proposició 3.1 $b\in a+F\Rightarrow a+F=b+F$.

DEMOSTRACIÓ: Per hipòtesi $\overrightarrow{ab}\in F$. Llavors

$$d\in a+F \ \Leftrightarrow\ \overrightarrow{ad}\in F\Leftrightarrow \overrightarrow{ab}+\overrightarrow{bd}=\overrightarrow{ab}-\overrightarrow{db}\in F\Leftrightarrow$$
$$\Leftrightarrow\ \overrightarrow{db}\in F\Leftrightarrow d\in b+F.\square$$

Corol·lari 3.2 $p, q \in a + F \Rightarrow \overrightarrow{pq} \in F$.

DEMOSTRACIÓ: Si $p \in a + F$, per (3.1) $a + F = p + F$; com que $q \in p + F$, $\overrightarrow{pq} \in F$.□

Proposició 3.3 *Donats* $a_1, \ldots, a_k \in A$ *existeix una varietat lineal mínima que conté* $a_1, \ldots, a_k$ *i que denominarem* varietat generada *per* $a_1, \ldots, a_k$.

Aquí "mínima" vol dir "continguda en qualsevol altra varietat que contingui també $a_1, \ldots, a_k$".

DEMOSTRACIÓ: Sigui $a_1 + F$ una varietat que conté $a_1, \ldots, a_k$ (n'existeix una: tot A). Per (3.2) $\overrightarrow{a_1 a_i} \in F \quad i = 2, \ldots, k$, d'on $\langle \overrightarrow{a_1 a_2}, \ldots, \overrightarrow{a_1 a_k} \rangle \subset F$.

La varietat

$$a_1 + \langle \overrightarrow{a_1 a_2}, \ldots, \overrightarrow{a_1 a_k} \rangle$$

conté a_1, $a_i = a_1 + \overrightarrow{a_1 a_i}$ $(i = 2, \ldots, k)$ i està continguda a $a_1 + F$; és, doncs, la varietat generada per $a_1, \ldots, a_k$. □

Per cada punt $a \in A$ passa una varietat lineal (i només una!) amb una direcció donada F. Dues varietats lineals amb la mateixa direcció direm que són paral·leles. Més en general, direm que dues varietats lineals $a + F$, $b + G$ són *paral·leles* si

$$F \subset G \quad \text{o} \quad G \subset F.$$

Nota:

Quan estudiàvem els espais vectorials o els grups ens ocupàvem també dels subconjunts que eren ells mateixos espais vectorials o grups amb les "mateixes" operacions del conjunt total: els subespais vectorials o els subgrups, respectivament. Què és ara un "subespai afí" d'un espai afí (A, E, φ)? Serà un espai afí (L, F, φ') on $L \subset A$, F sigui un subespai de E i

$$\varphi' : L \times L \longrightarrow F$$

sigui la restricció de φ. És a dir, s'haurà de complir

$$\forall p, q \in L, \qquad \overrightarrow{pq} = \varphi(p, q) \in F$$

i, llavors, $\varphi'(p, q) = \varphi(p, q)$. En particular, si $p \in L$, tot altre $q \in L$ és de la forma $q = p + \overrightarrow{pq} \in p + F$, i $L \subset p + F$. Recíprocament, tot $q = p + u \in p + F$ és tal que $q = \varphi_p^{-1}(u)$. Com que, clarament, $q = \varphi_p^{-1}(u) = \varphi'^{-1}_p(u)$, resulta que $q \in L$ i, per tant, $p + F \subset L$.

Hem provat d'aquesta manera que les varietats lineals són precisament els "subespais afins" d'un espai afí.

IX.4 Intersecció i suma de varietats lineals

Proposició 4.1 *Dues varietats lineals $a+F$, $b+G$ es tallen si i només si*

$$\overrightarrow{ab} \in F+G.$$

DEMOSTRACIÓ: Suposem primer que es tallen i sigui c un punt de la intersecció $(a+F)\cap(b+G)$. Llavors $\overrightarrow{ac}\in F$ i $\overrightarrow{bc}\in G$, d'on $\overrightarrow{ab}=\overrightarrow{ac}-\overrightarrow{bc}\in F+G$.

Suposem ara que $\overrightarrow{ab}=u+v$ amb $u\in F$ i $v\in G$. Sigui c un punt tal que $\overrightarrow{ac}=u$ (en particular, $c\in a+F$). Llavors $v=\overrightarrow{ab}-u=\overrightarrow{ab}-\overrightarrow{ac}=\overrightarrow{cb}$ i, per tant, $\overrightarrow{bc}=-v\in G$, d'on $c\in b+G$. El punt c és, doncs, un punt comú a les dues varietats.□

Corol·lari 4.2 *Dues varietats paral·leles o no es tallen o una està continguda en l'altra.*

DEMOSTRACIÓ: Sigui $a+F$, $b+G$ amb $F\subset G$.

$$(a+F)\cap(b+G)\neq\emptyset \Leftrightarrow \overrightarrow{ab}\in F+G=G \Leftrightarrow a\in b+G \Leftrightarrow$$

$$\Leftrightarrow a+F\subset a+G=b+G.\square$$

Proposició 4.3 *Si $a+F$ i $b+G$ tenen un punt en comú, c,*

$$(a+F)\cap(b+G)=c+(F\cap G).$$

DEMOSTRACIÓ: Per (3.1) $a+F=c+F$ i $b+G=c+G$. Llavors

$$x\in(c+F)\cap(c+G)\Leftrightarrow \overrightarrow{cx}\in F \text{ i } \overrightarrow{cx}\in G\Leftrightarrow \overrightarrow{cx}\in F\cap G\Leftrightarrow$$

$$\Leftrightarrow x\in c+(F\cap G).\square$$

Es diu que dues varietats lineals es *creuen* si no són paral·leles ni es tallen.

Acabem de veure que la intersecció de dues varietats lineals, si no és buida, és una varietat lineal (4.3). La unió de dues varietats lineals, en canvi, no és en general una varietat.

Exemple:

Siguin u, v dos vectors linealment independents.

Els punts $c=a+u$ i $d=a+v$ són de la unió $L=(a+\langle v\rangle)\cup(a+\langle u\rangle)$. Si L fos una varietat lineal, el vector $\overrightarrow{cd}=v-u$ seria de la seva direcció (3.2) i el punt $p=a+(v-u)$ seria de L. Ara bé, el punt p no és de $a+\langle v\rangle$ ni de $a+\langle u\rangle$ i, per tant, no pot ésser de L.

Donades dues varietats lineals $a+F$ i $b+G$, quina és la varietat lineal "mínima"que les conté? Suposem que $a+H$ conté $a+F$ i $b+G$. Llavors

$$H \supset F+G+\langle \overrightarrow{ab} \rangle.$$

En efecte, $a, b \in a+H$, d'on $\overrightarrow{ab} \in H$. A més a més,

$$\begin{array}{rcl} u \in F & \Rightarrow & p = a+u \in a+F \subset a+H \Rightarrow u = \overrightarrow{ap} \in H \\ v \in G & \Rightarrow & q = b+v \in b+G \subset a+H \Rightarrow \overrightarrow{aq} \in H \\ & \Rightarrow & v = \overrightarrow{bq} = \overrightarrow{aq} - \overrightarrow{ab} \in H. \end{array}$$

Per altra banda, la varietat

$$a+F+G+\langle \overrightarrow{ab} \rangle$$

conté clarament $a+F$ i $b+G$. Aquesta és, per tant, la varietat lineal "mínima" que conté $a+F$ i $b+G$ i en direm *varietat suma*: $(a+F)+(b+G)$.

Proposició 4.4 (Fórmules de Grassmann) *Siguin $L_1 = a+F$ i $L_2 = b+G$ dues varietats lineals i sigui L_1+L_2 la seva suma. Aleshores, si $L_1 \cap L_2 \neq \emptyset$,*

$$dim\ (L_1+L_2) = dim\ L_1 + dim\ L_2 - dim\ (L_1 \cap L_2),$$

i si $L_1 \cap L_2 = \emptyset$,

$$dim\ (L_1+L_2) = dim\ L_1 + dim\ L_2 - dim\ (F \cap G) + 1.$$

DEMOSTRACIÓ: Clarament $\dim\ (L_1+L_2) = \dim\ (F+G+\langle \overrightarrow{ab} \rangle)$. Per (4.1), si $L_1 \cap L_2 \neq \emptyset$, $\overrightarrow{ab} \in F+G$ i

$$\dim\ (L_1+L_2) = \dim\ (F+G) = \dim\ F + \dim\ G - \dim\ (F \cap G).$$

En aquest cas, $L_1 \cap L_2$ és una varietat lineal de direcció $F \cap G$ i $\dim\ (F \cap G) = \dim\ (L_1 \cap L_2)$. Llavors

$$\dim\ (L_1+L_2) = \dim\ L_1 + \dim\ L_2 - \dim\ (L_1 \cap L_2).$$

Si $L_1 \cap L_2 = \emptyset$, per (4.1), $\overrightarrow{ab} \notin F+G$ i

$$\begin{array}{rcl} \dim\ (L_1+L_2) & = & \dim\ (F+G)+1 = \\ & = & \dim\ F + \dim\ G - \dim\ (F \cap G) + 1 = \\ & = & \dim\ L_1 + \dim\ L_2 - \dim\ (F \cap G) + 1. \square \end{array}$$

Exemple:

Sigui A un espai afí de dimensió 3. Siguin

$$r = a + F \qquad \pi = b + G$$

una recta i un pla de A: dim $F = 1$, dim $G = 2$. Si $r \cap \pi = \emptyset$,

$$\dim (r + \pi) = 4 - \dim (F \cap G).$$

Però aquesta dimensió ha d'ésser més petita que 3; per tant, dim $(F \cap G) = 1$, d'on $F \subset G$ i r i π són paral·leles.

Si $r \cap \pi \neq \emptyset$,

$$\dim (r + \pi) = 3 - \dim (r \cap \pi).$$

Per tant,

ó bé $\dim (r + \pi) = 3 \Rightarrow \dim (r \cap \pi) = 0 \quad \Rightarrow r \cap \pi =$ un punt,
ó bé $\dim (r + \pi) = 2 \Rightarrow \dim (r \cap \pi) = 1 \quad \Rightarrow r \subset \pi.$

En aquest segon cas ($r \subset \pi$), r i π són paral·leles. Així doncs, en un espai afí de dimensió 3 una recta i un pla no paral·lels sempre es tallen en un punt.

Exercici:

Demostreu que en un espai afí de dimensió 2 dues rectes no paral·leles sempre es tallen en un punt.

IX.5 Dependència lineal de punts

Proposició 5.1 *Sigui (A, E) un espai afí i siguin $a_1, \ldots, a_k \in A$. Les següents condicions són equivalents:*

a) Els vectors $\overrightarrow{a_1a_2}, \ldots, \overrightarrow{a_1a_k}$ de E són linealment independents.

b) Fixat qualsevol i, els vectors $\{\overrightarrow{a_ia_h},\ h \neq i\}$ són linealment independents.

c) Per a tot $p \in A$

$$\left.\begin{array}{c} \lambda^1\overrightarrow{pa_1} + \ldots + \lambda^k\overrightarrow{pa_k} = 0 \\ \lambda^1 + \ldots + \lambda^k = 0 \end{array}\right\} \Rightarrow \lambda^1 = \ldots = \lambda^k = 0.$$

DEMOSTRACIÓ: a) $\Rightarrow$ b). En efecte, de

$$\sum_{h \neq i} \lambda^h \overrightarrow{a_ia_h} = \vec{0}$$

resulta

$$\sum_{h\neq i}\lambda^h(\overrightarrow{a_1a_h}-\overrightarrow{a_1a_i})=\sum_{h\neq i}\lambda^h\overrightarrow{a_1a_h}-\left(\sum_{h\neq i}\lambda^h\right)\overrightarrow{a_1a_i}=\vec{0}.$$

a) ens diu aleshores que $\lambda^h=0\ \forall h\neq i,1$. Però $\sum_{h\neq i}\lambda^h=0$ implica també $\lambda^1=0$.

b) $\Rightarrow$ a). En efecte, a) coincideix amb b) per a $i=1$.

a) $\Rightarrow$ c). En efecte, de

$$\begin{cases}\lambda^1\overrightarrow{pa_1}+\ldots+\lambda^k\overrightarrow{pa_k}=\vec{0}\\ \lambda^1+\ldots+\lambda^k=0\end{cases}$$

es dedueix que

$$\begin{aligned}&(-\lambda^2-\ldots-\lambda^k)\overrightarrow{pa_1}+\lambda^2\overrightarrow{pa_2}+\ldots+\lambda^k\overrightarrow{pa_k}=\\&=\lambda^2(\overrightarrow{pa_2}-\overrightarrow{pa_1})+\ldots+\lambda^k(\overrightarrow{pa_k}-\overrightarrow{pa_1})=\\&=\lambda^2\overrightarrow{a_1a_2}+\ldots+\lambda^k\overrightarrow{a_1a_k}=\vec{0}.\end{aligned}$$

a) ens diu llavors que $\lambda^2=\ldots=\lambda^k=0$ i, també, que $\lambda^1=-\lambda^2-\ldots-\lambda^k=0$.

c) $\Rightarrow$ a). En efecte, de

$$\sum_{h\geq 2}\lambda^h\overrightarrow{a_1a_h}=\vec{0}$$

s'obté

$$\sum_{h\geq 2}\lambda^h(\overrightarrow{pa_h}-\overrightarrow{pa_1})=\left(-\sum_{h\geq 2}\lambda^h\right)\overrightarrow{pa_1}+\sum_{h\geq 2}\lambda^h\overrightarrow{pa_h}=\vec{0}.$$

En aquesta expressió la suma dels coeficients és 0 i, per tant, c) assegura que $\lambda^h=0$, $\forall h\geq 2$. □

Direm que els punts $a_1,\ldots,a_k\in A$ són *linealment independents* si compleixen qualsevol de les condicions de (5.1).

Observacions:

1. Dos punts són linealment independents si i només si són diferents.
2. Si dim $A=n$, el nombre màxim de punts linealment independents és $n+1$.

Proposició 5.2 *Donats $a_1,\ldots,a_k$ punts linealment independents d'un espai afí A de dimensió n, existeixen punts $a_{k+1},\ldots,a_{n+1}$ de A tals que $a_1,\ldots,a_k,\ldots,a_{n+1}$ són linealment independents.*

DEMOSTRACIÓ: Si $a_1,\ldots,a_k$ són linealment independents, llavors també ho són els vectors $\overrightarrow{a_1a_2},\ldots,\overrightarrow{a_1a_k}$. Siguin $u_k,\ldots,u_n$ vectors tals que

$$\overrightarrow{a_1a_2},\ldots,\overrightarrow{a_1a_k},u_k,\ldots,u_n$$

sigui una base de l'espai vectorial associat a A. Considerem punts a_i tals que $\overrightarrow{a_1a_i}=u_{i-1}$, $i=k+1,\ldots,n+1$. Els punts $a_1,\ldots,a_k,\ldots,a_{n+1}$ són clarament linealment independents. □

IX.6 Coordenades baricèntriques

Sigui (A,E) un espai afí de dimensió n i siguin $p_0,p_1,\ldots,p_n$ punts linealment independents de A. Donat un $x\in A$, els punts $p_0,p_1,\ldots,p_n,x$ no poden ser linealment independents i, per tant, existeixen $p\in A$ i elements $k,k^0,\ldots,k^n\in K$, no tots zero, tals que

$$\left.\begin{array}{c} k\overrightarrow{px}+k^0\overrightarrow{pp_0}+\ldots+k^n\overrightarrow{pp_n}=\overrightarrow{0}\\ k+k^0+\ldots+k^n=0\end{array}\right\}.$$

Si $k=0$, aquestes igualtats implicarien que $p_0,\ldots,p_n$ fossin linealment dependents, i no ho són per hipòtesi. Per tant, $k\neq 0$ i, posant $x^i=-k^i/k$, tenim

$$\left.\begin{array}{c}\overrightarrow{px}=x^0\overrightarrow{pp_0}+\ldots+x^n\overrightarrow{pp_n}\\ x^0+\ldots+x^n=1\end{array}\right\}.$$

Anem a veure que aquests $x^0,\ldots,x^n$ no depenen del punt p.

Proposició 6.1 *Siguin $p_0,p_1,\ldots,p_k$ punts d'un espai afí (A,E) de dimensió n. Suposem que, donat $x\in A$, existeixen $p\in A$ i $x^0,\ldots,x^k\in K$ tals que*

$$\left.\begin{array}{c}\overrightarrow{px}=x^0\overrightarrow{pp_0}+\ldots+x^k\overrightarrow{pp_k}\\ x^0+\ldots+x^k=1\end{array}\right\}.$$

Aleshores

i) Aquestes expressions també són certes si substituïm p per qualsevol altre $q\in A$.

ii) Els valors $x^0,\ldots,x^k$ estan unívocament determinats, sempre que $p_0,\ldots,p_k$ siguin linealment independents.

DEMOSTRACIÓ:

$$\begin{aligned}\overrightarrow{qx}=\overrightarrow{qp}+\overrightarrow{px} &= (x^0+\ldots+x^k)\overrightarrow{qp}+x^0\overrightarrow{pp_0}+\ldots+x^k\overrightarrow{pp_k}=\\ &= x^0(\overrightarrow{qp}+\overrightarrow{pp_0})+\ldots+x^k(\overrightarrow{qp}+\overrightarrow{pp_k})=\\ &= x^0\overrightarrow{qp_0}+\ldots+x^k\overrightarrow{qp_k}.\end{aligned}$$

Això demostra i). Per provar ii) apliquem i) al cas $q = p_0$; obtenim

$$\overrightarrow{p_0x} = x^1\overrightarrow{p_0p_1} + \ldots + x^k\overrightarrow{p_0p_k}.$$

$\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_k}$ són vectors linealment independents i, per tant, $x^1, \ldots, x^k$ estan unívocament determinats. D'aquí que $x^0 = 1 - x^1 - \ldots - x^k$ també estigui determinat. □

Anomenarem *sistema de referència baricèntric* o *sistema de coordenades baricèntric* d'un espai afí (A, E) de dimensió n tot conjunt de $(n+1)$ punts linealment independents $\{p_0, \ldots, p_n\}$. Anomenarem *coordenades baricèntriques* d'un punt $x \in A$ en el sistema $\{p_0, \ldots, p_n\}$ la $(n+1)$-pla $(x^0, \ldots, x^n) \in K^{n+1}$ tal que

$$\left.\begin{array}{c} \overrightarrow{px} = x^0\overrightarrow{pp_0} + \ldots + x^n\overrightarrow{pp_n} \\ x^0 + \ldots + x^n = 1 \end{array}\right\}.$$

Per (6.1) aquesta $(n+1)$-pla és independent del punt $p \in A$ i està ben determinada.

Notació:

Donats $p_0, \ldots, p_k$, si

$$\left.\begin{array}{c} \overrightarrow{px} = x^0\overrightarrow{pp_0} + \ldots + x^k\overrightarrow{pp_k} \\ x^0 + \ldots + x^k = 1 \end{array}\right\},$$

escriurem

$$x = x^0p_0 + \ldots + x^kp_k.$$

Tinguem en compte, però, que aquesta expressió no té cap sentit si la suma dels coeficients no és 1.

Es diu *baricentre* de m punts $a_1, \ldots, a_m \in A$ un punt $b = b^1a_1 + \ldots + b^ma_m$, $\sum b^i = 1$, tal que $b^1 = \ldots = b^m$. Perquè existeixi el baricentre de m punts ha d'haver-hi un element $d \in K$ tal que $\overbrace{d + \ldots + d}^{m} = 1$. Designem també per m l'element de K que resulta de sumar m vegades $1 \in K$:

$$m = \overbrace{1 + \ldots + 1}^{m} \in K.$$

Aleshores

$$1 = \overbrace{d + \ldots + d}^{m} = d(\overbrace{1 + \ldots + 1}^{m}) = dm.$$

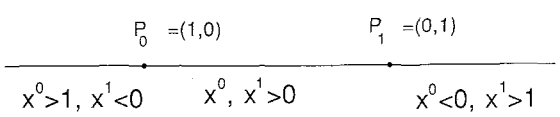
d existeix només si $m \neq 0$ a K i llavors $d = m^{-1}$ (escriurem també $d = \frac{1}{m}$).

Nota:

El nucli de l'aplicació $f : \mathbf{Z} \longrightarrow K$ tal que $f(m) = m$ si $m > 0$, $f(m) = -f(-m)$ si $m < 0$ i $f(0) = 0$ és un ideal (p), $p \geq 0$. En aquesta situació es diu que K *és un cos de característica* p. Si $p \neq 0$, p és el més petit enter positiu que és 0 a K. Ha d'ésser un nombre primer, ja que en cas contrari, si $p = r \cdot m$, tindríem que a K $0 = r \cdot m$ amb $r \neq 0$, $m \neq 0$. Els cossos $\mathbf{Q}$, $\mathbf{R}$ i $\mathbf{C}$ són de característica 0; per a p primer, el cos $\mathbf{Z}/(p)$ és de característica p.

Exemples:

1. En el sistema de referència baricèntric $\{p_0, \ldots, p_n\}$ les coordenades baricèntriques de p_0 són $(1,0,\ldots,0)$; les de p_1 són $(0,1,0,\ldots,0)$;...; les de p_n són $(0,\ldots,0,1)$.
2. Sigui $(\mathbf{R},\mathbf{R})$ la recta afí real. Sigui $\{p_0,p_1\}$ una referència baricèntrica de $\mathbf{R}$. Sigui x un punt de coordenades (x^0,x^1) (amb $x^0 + x^1 = 1$) en aquest sistema.
   Llavors $\overrightarrow{px} = x^0\overrightarrow{pp_0} + x^1\overrightarrow{pp_1}$, $\forall p$, i, en particular, $\overrightarrow{p_0x} = x^1\overrightarrow{p_0p_1}$.



   Això ens permet conèixer la posició de x respecte als punts p_0, p_1, segons que les seves coordenades siguin positives o negatives (veure figura). En particular, el *segment* determinat per p_0 i p_1 és el conjunt

$$\overline{p_0p_1} = \{x \in A \mid x = x^0p_0 + x^1p_1,\ x^0 + x^1 = 1,\ 0 \leq x^0, x^1 \leq 1\}.$$

   Es diu *punt mig* del segment $\overline{p_0p_1}$ el punt $\frac{1}{2}p_0 + \frac{1}{2}p_1$, és a dir, el baricentre de $\{p_0,p_1\}$.
3. Sigui $(\mathbf{R}^2,\mathbf{R}^2)$ el pla afí real i $\{p_0,p_1,p_2\}$ un sistema de referència baricèntric de $\mathbf{R}^2$. Sigui x un punt, i (x^0,x^1,x^2) les seves coordenades en aquest sistema:

$$\overrightarrow{px} = x^0\overrightarrow{pp_0} + x^1\overrightarrow{pp_1} + x^2\overrightarrow{pp_2} \qquad x^0 + x^1 + x^2 = 1 \qquad \forall p.$$

Si x^0=1, $x^2 = -x^1$ i

$$\overrightarrow{px} = \overrightarrow{pp_0} + x^1(\overrightarrow{pp_1} - \overrightarrow{pp_2}) = \overrightarrow{pp_0} + x^1\overrightarrow{p_2p_1} \qquad \forall p.$$

Per tant, $\overrightarrow{p_0x} = x^1\overrightarrow{p_2p_1}$; és a dir, x està sobre la recta $p_0 + <\overrightarrow{p_2p_1}>$ (veure figura).

Si $x^0 \neq 1$, resulta que, per a tot p,

$$\overrightarrow{px} = x^0\overrightarrow{pp_0} + (1-x^0)\left(\frac{x^1}{1-x^0}\overrightarrow{pp_1} + \frac{x^2}{1-x^0}\overrightarrow{pp_2}\right).$$

Sigui q el punt tal que

$$\overrightarrow{pq} = \frac{x^1}{1-x^0}\overrightarrow{pp_1} + \frac{x^2}{1-x^0}\overrightarrow{pp_2}.$$

La suma de coeficients d'aquesta expressió és 1 i, per tant, de (6.1) resulta que

$$\overrightarrow{p_1q} = \frac{x^2}{1-x^0}\overrightarrow{p_1p_2}$$

i q és un punt de la recta $p_1 + \langle\overrightarrow{p_1p_2}\rangle$. Si apliquem ara els resultats de l'exemple anterior a

$$\overrightarrow{px} = x^0\overrightarrow{pp_0} + (1-x^0)\overrightarrow{pq},$$

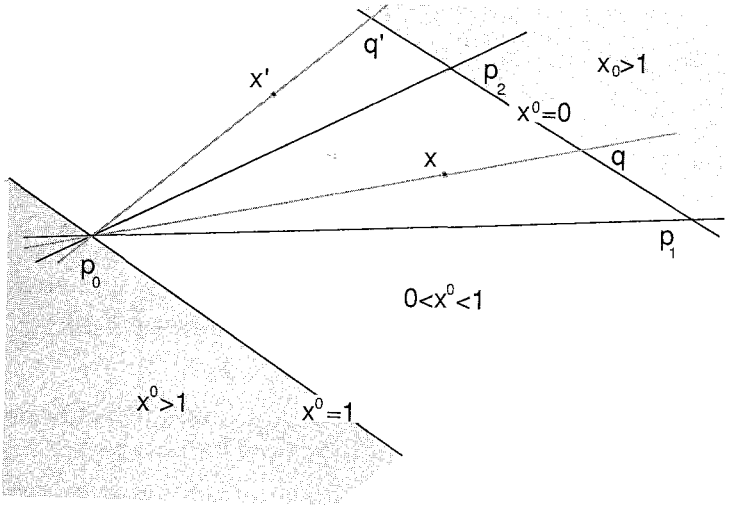
podem deduir la posició de x segons que $x^0 < 0$, $0 \leq x^0 \leq 1$ o $x^0 > 1$, tal com està indicat a la figura. Naturalment, podem obtenir resultats anàlegs per a les dues altres coordenades i, d'aquesta forma, determinar la regió del pla on és el punt x, segons els signes de les seves coordenades.

Exercici:

Generalitzeu l'estudi fet a l'exemple 2 a l'espai afí $\mathbf{R}^3$, estudiant la regió on és un punt x segons els signes de les seves coordenades baricèntriques respecte a un sistema de referència $\{p_0, p_1, p_2, p_3\}$.

Acabarem aquest apartat estudiant com canvien les coordenades baricèntriques d'un punt en canviar el sistema de referència. Siguin $\{p_0, \ldots, p_n\}$ i $\{q_0, \ldots, q_n\}$ dos sistemes de referència baricèntrics d'un espai afí A. Siguin $(q_i^0, \ldots, q_i^n)$ les coordenades baricèntriques de q_i en el sistema $\{p_0, \ldots, p_n\}$, $i = 0, \ldots, n$. Donat $x \in A$, indiquem per $(x^0, \ldots, x^n)$ i $(\bar{x}^0, \ldots, \bar{x}^n)$ les seves coordenades baricèntriques en els sistemes $\{p_0, \ldots, p_n\}$ i $\{q_0, \ldots, q_n\}$ respectivament. Aleshores, per a cada p,

$$\begin{aligned} \overrightarrow{px} &= \textstyle\sum_{i=0}^n \bar{x}^i\overrightarrow{pq_i} = \sum_{i=0}^n \bar{x}^i\left(\sum_{j=0}^n q_i^j\overrightarrow{pp_j}\right) = \\ &= \textstyle\sum_{j=0}^n \left(\sum_{i=0}^n \bar{x}^i q_i^j\right)\overrightarrow{pp_j}, \end{aligned}$$



amb suma de coeficients

$$\sum_{j=0}^{n}\left(\sum_{i=0}^{n}\bar{x}^i q_i^j\right)=\sum_{i=0}^{n}\bar{x}^i\left(\sum_{j=0}^{n}q_i^j\right)=\sum_{i=0}^{n}\bar{x}^i=1.$$

Per tant,

$$\sum_{i=0}^{n}q_i^j\bar{x}^i=x^j \qquad j=0,\ldots,n.$$

Aquestes $(n+1)$ expressions equivalen a la igualtat matricial

$$Q\bar{x}=x,$$

on $Q=\left(q_i^j\right)$ és la matriu que té per columnes les coordenades baricèntriques de $q_0,\ldots,q_n$, i $\bar{x}$, x són les matrius d'una columna formades per les coordenades de x en els sistemes $\{q_0,\ldots,q_n\}$ i $\{p_0,\ldots,p_n\}$, respectivament.

Observació:

La matriu Q és invertible, ja que

$$\det Q=\begin{vmatrix} q_0^0 & \cdots & q_n^0 \\ \vdots & & \vdots \\ q_0^n & \cdots & q_n^n \end{vmatrix}=$$

(sumant totes les files a la primera)

$$= \begin{vmatrix} 1 & \dots & 1 \\ q_0^1 & \dots & q_n^1 \\ \vdots & & \vdots \\ q_0^n & \dots & q_n^n \end{vmatrix} = \begin{vmatrix} 1 & 0 & \dots & 0 \\ q_0^1 & q_1^1 - q_0^1 & \dots & q_n^1 - q_0^1 \\ \vdots & \vdots & & \vdots \\ q_0^n & q_1^n - q_0^n & \dots & q_n^n - q_0^n \end{vmatrix} =$$

$$= \det_{\overrightarrow{(p_0 p_i)}}(\overrightarrow{q_0 q_1}, \dots, \overrightarrow{q_0 q_n}) \neq 0.$$

IX.7 Equacions d'una varietat en coordenades baricèntriques

Sigui (A, E) un espai afí de dimensió n i $\{p_0, \dots, p_n\}$ un sistema de referència baricèntric de A. Donats k punts $a_1, \dots, a_k$, les seves coordenades baricèntriques $(a_i^0, \dots, a_i^n)$ compleixen

$$\operatorname{rang} \begin{pmatrix} a_1^0 & \dots & a_k^0 \\ \vdots & & \vdots \\ a_1^n & \dots & a_k^n \end{pmatrix} = \operatorname{rang} \begin{pmatrix} 1 & \dots & 1 \\ a_1^1 & \dots & a_k^1 \\ \vdots & & \vdots \\ a_1^n & \dots & a_k^n \end{pmatrix} =$$

$$= \operatorname{rang} \begin{pmatrix} 1 & 0 & \dots & 0 \\ a_1^1 & a_2^1 - a_1^1 & \dots & a_k^1 - a_1^1 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n - a_1^n & \dots & a_k^n - a_1^n \end{pmatrix} =$$

$$= \dim \langle \overrightarrow{a_1 a_2}, \dots, \overrightarrow{a_1 a_k} \rangle + 1.$$

En particular, $a_1, \dots, a_k$ són linealment independents si i només si aquest rang és k.

Designem per L la varietat lineal generada per $a_1, \dots, a_k$ (3.3). Tenim

$$L = \{x \in A;\ \overrightarrow{a_1 x} = \lambda^2 \overrightarrow{a_1 a_2} + \dots + \lambda^k \overrightarrow{a_1 a_k}\}.$$

Ara bé,

$$\begin{aligned} \overrightarrow{a_1 x} &= \lambda^2 \overrightarrow{a_1 a_2} + \dots + \lambda^k \overrightarrow{a_1 a_k} = \\ &= (1 - \lambda^2 - \dots - \lambda^k) \overrightarrow{a_1 a_1} + \lambda^2 \overrightarrow{a_1 a_2} + \dots + \lambda^k \overrightarrow{a_1 a_k} \end{aligned}$$

i, per (6.1),

$$\overrightarrow{px} = \lambda^1 \overrightarrow{pa_1} + \dots + \lambda^k \overrightarrow{pa_k} \qquad \forall p,$$

on $\lambda^1 = 1 - \lambda^2 - \ldots - \lambda^k$. És a dir,

$$L = \{x \in A \mid x = \lambda^1 a_1 + \ldots + \lambda^k a_k,\ \lambda^1 + \ldots + \lambda^k = 1\}.$$

D'aquí resulta immediatament que

$$x^i = \lambda^1 a_1^i + \ldots + \lambda^k a_k^i \qquad i = 0, \ldots, n,$$

on (x^i), $(a_1^i), \ldots$, (a_k^i) són les coordenades baricèntriques de $x, a_1, \ldots, a_k$ respectivament en un cert sistema de referència $\{p_0, \ldots, p_n\}$.

Suposem que $a_1, \ldots, a_k$ són linealment independents i, per tant, dim $L = k - 1$ (3.3). Llavors $x \in L$ si i només si $x, a_1, \ldots, a_k$ són linealment dependents, és a dir, si

$$k = \operatorname{rang} \begin{pmatrix} a_1^0 & \ldots & a_k^0 \\ \vdots & & \vdots \\ a_1^n & \ldots & a_k^n \end{pmatrix} = \operatorname{rang} \begin{pmatrix} a_1^0 & \ldots & a_k^0 & x^0 \\ \vdots & & \vdots & \vdots \\ a_1^n & \ldots & a_k^n & x^n \end{pmatrix}.$$

Fixem un menor d'ordre k de la primera matriu, que tingui determinant diferent de zero. Per comoditat suposarem que està format per les k primeres files. Aleshores una $(n+1)$–pla $(x^0, \ldots, x^n)$ és el conjunt de coordenades baricèntriques d'un punt $x \in L$ si i només si compleix:

$$\begin{vmatrix} a_1^0 & \ldots & a_k^0 & x^0 \\ \vdots & & \vdots & \vdots \\ a_1^{k-1} & \ldots & a_k^{k-1} & x^{k-1} \\ a_1^i & \ldots & a_k^i & x^i \end{vmatrix} = 0 \qquad i = k, \ldots, n$$

$$\textstyle\sum_{j=0}^{n} x^j = 1.$$

Aquestes són les *equacions baricèntriques* de la varietat lineal L.

Exemple:

Les equacions de l'hiperplà determinat pels punts del sistema de referència $p_0, \ldots, p_n$ llevat d'un d'ells, p_i, són

$$\left.\begin{aligned} x^i &= 0 \\ x^0 + \ldots + x^n &= 1 \end{aligned}\right\}.$$

IX.8 Coordenades cartesianes

Fixem un punt p de l'espai afí (A, E). L'aplicació

$$\begin{array}{rccc} \varphi_p : & A & \longrightarrow & E \\ & q & \longmapsto & \overrightarrow{pq} \end{array}$$

sabem que és bijectiva i, per tant, fixada una base de E $e_1, \ldots, e_n$, el punt q queda completament determinat per les coordenades de $\overrightarrow{pq}$ en aquesta base

$$\overrightarrow{pq} = q^1 e_1 + \ldots + q^n e_n.$$

Anomenarem *referència cartesiana* o *sistema de coordenades cartesianes* de (A, E) el conjunt $\{p; e_1, \ldots, e_n\}$ format per un punt $p \in A$ i una base $e_1, \ldots, e_n$ de E. Anomenarem *coordenades cartesianes* d'un punt q en aquest sistema les coordenades de $\overrightarrow{pq}$ en la base $e_1, \ldots, e_n$.

Exemple:

Les coordenades de p en el sistema $\{p; e_1, \ldots, e_n\}$ són $(0, \ldots, 0)$.

Observació:

Les coordenades d'un vector $\overrightarrow{ab}$ en la base $e_1, \ldots, e_n$, són les coordenades de $\overrightarrow{ab} = \overrightarrow{pb} - \overrightarrow{pa}$ i, per tant, són la diferència de les coordenades cartesianes de b i a en el sistema $\{p; e_1, \ldots, e_n\}$. Aquest fet justifica la notació introduïda al §3:

$$\overrightarrow{ab} = b - a \qquad b = a + \overrightarrow{ab}.$$

Considerem ara dos sistemes de referència $\{p; e_1, \ldots, e_n\}$, $\{q; v_1, \ldots, v_n\}$ i suposem conegut el segon en funció del primer:

$$\overrightarrow{pq} = q^1 e_1 + \ldots + q^n e_n \qquad v_i = \sum_{j=1}^{n} v_i^j e_j \qquad i = 1, \ldots, n.$$

Donat $x \in A$, volem estudiar la relació entre les seves coordenades en un sistema i en l'altre.

$$\overrightarrow{px} = x^1 e_1 + \ldots + x^n e_n \qquad \overrightarrow{qx} = \bar{x}^1 v_1 + \ldots + \bar{x}^n v_n.$$

Tenim

$$\left.\begin{array}{rcl} \overrightarrow{qx} & = & \sum_{i=1}^n \bar{x}^i v_i = \sum_{i=1}^n \bar{x}^i \left(\sum_{j=1}^n v_i^j e_j\right) = \\ & = & \sum_{j=1}^n \left(\sum_{i=1}^n v_i^j \bar{x}^i\right) e_j \\ \overrightarrow{qx} & = & \overrightarrow{px} - \overrightarrow{pq} = \sum_{j=1}^n (x^j - q^j) e_j \end{array}\right\},$$

d'on $x^j = q^j + \sum_{i=1}^n v_i^j \bar{x}^i$, $j = 1, \ldots, n$.

Aquestes equacions es poden expressar matricialment així: sigui $V = \left(v_i^j\right)$ i x, $\bar{x}$, q les matrius d'una columna formades per les coordenades de x, $\bar{x}$ i q, respectivament. Aleshores

$$x = q + V\bar{x};$$

o també

$$\begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} V & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{x} \\ 1 \end{pmatrix},$$

amb la notació òbvia.

IX.9 Equacions d'una varietat en coordenades cartesianes

Sigui $L = a + \langle v_1, \ldots, v_k \rangle$ una varietat lineal de dimensió k de l'espai afí (A, E). Fixat un sistema de referència cartesià $\{p; e_1, \ldots, e_n\}$, ens proposem trobar les condicions que han de complir les coordenades d'un punt x perquè sigui de la varietat L. Tenim

$$x \in a + \langle v_1, \ldots, v_k \rangle \Leftrightarrow \overrightarrow{ax} \in \langle v_1, \ldots, v_k \rangle \Leftrightarrow$$

$$\operatorname{rang} \begin{pmatrix} x^1 - a^1 & v_1^1 & \ldots & v_k^1 \\ \vdots & \vdots & & \vdots \\ x^n - a^n & v_1^n & \ldots & v_k^n \end{pmatrix} = \operatorname{rang} \begin{pmatrix} v_1^1 & \ldots & v_k^1 \\ \vdots & & \vdots \\ v_1^n & \ldots & v_k^n \end{pmatrix} = k.$$

De la segona matriu, escollim un menor d'ordre k amb determinant diferent de zero. Per simplificar la notació suposarem que està format per les k primeres files:

$$\begin{vmatrix} v_1^1 & \ldots & v_k^1 \\ \vdots & & \vdots \\ v_1^k & \ldots & v_k^k \end{vmatrix} \neq 0.$$

Aleshores, les condicions que han de complir les coordenades $(x^1, \ldots, x^n)$ de $x \in L$ són

$$\begin{vmatrix} x^1 - a^1 & v_1^1 & \ldots & v_k^1 \\ \vdots & \vdots & & \vdots \\ x^k - a^k & v_1^k & \ldots & v_k^k \\ x^i - a^i & v_1^i & \ldots & v_k^i \end{vmatrix} = 0 \qquad i = k+1, \ldots, n.$$

Observem que aquests determinants formen un sistema de $n - k$ equacions amb n incògnites.

Recíprocament, sigui

$$\left.\begin{array}{cccc} a_1^1 x^1 + & \ldots & + a_n^1 x^n & = b^1 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^d x^1 + & \ldots & + a_n^d x^n & = b^d \end{array}\right\}$$

un sistema d'equacions lineals qualsevol. Les solucions del sistema homogeni associat, interpretades com a coordenades de vectors de E en la base $e_1, \ldots, e_n$, formen un subespai vectorial F de E. Sigui $(a^1, \ldots, a^n)$ una solució particular del sistema donat i sigui a el punt amb aquestes coordenades respecte al sistema de referència $\{p; e_1, \ldots, e_n\}$. La solució general del sistema és la suma de $(a^1, \ldots, a^n)$ i les solucions del sistema homogeni; és a dir, el conjunt de coordenades dels punts de la varietat lineal

$$a + F.$$

En aquestes circumstàncies direm que el sistema donat són les *equacions de la varietat* $a + F$ en el sistema de referència cartesià $\{p; e_1, \ldots, e_n\}$. Observem que la dimensió d'aquesta varietat és $n - \text{rang}\left(a_i^j\right)$.

Exemples:

1. Una equació $a_1x^1 + \ldots + a_nx^n = b$ no trivial (és a dir, amb no tots els coeficients a_i zero) representa un hiperplà.

   En general, si

   $$\left.\begin{array}{cccc} a_1^1x^1 + & \ldots & +a_n^1x^n & = & b^1 \\ \ldots & \ldots & \ldots & & \ldots \\ a_1^dx^1 + & \ldots & +a_n^dx^n & = & b^d \end{array}\right\}$$

   són les equacions d'una varietat L, cada una de les equacions representa un hiperplà H_i i la varietat L resulta ésser la intersecció d'aquests d hiperplans. L'equació de qualsevol altre hiperplà que contingui L serà tal que, afegida al sistema, aquest tindrà les mateixes solucions; és a dir, serà una combinació lineal de les equacions del sistema:

   $$\lambda^1(a_1^1x^1 + \ldots + a_n^1x^n) + \ldots + \lambda^d(a_1^dx^1 + \ldots + a_n^dx^n) =$$
   $$= \lambda^1b^1 + \ldots + \lambda^db^d.$$

   Aquest conjunt d'hiperplans es diu el *feix d'hiperplans* que passen per L.

2. Les equacions d'una recta $a + \langle v \rangle$ són

   $$\begin{vmatrix} x^i - a^i & v^i \\ x^j - a^j & v^j \end{vmatrix} = 0 \qquad \forall i, j.$$

   És a dir, $(x^i - a^i)v^j = (x^j - a^j)v^i \ \forall i, j$. Aquestes equacions les podem escriure així:

   $$\frac{x^1 - a^1}{v^1} = \ldots = \frac{x^n - a^n}{v^n},$$

   amb el conveni que si algun $v^j = 0$ el quocient corresponent té numerador zero: $x^j - a^j = 0$.

L'estudi dels sistemes d'equacions de dues varietats permet estudiar fàcilment qüestions com intersecció o paral·lelisme. Anem a veure dos casos de paral·lelisme particularment senzills.

I. Dos hiperplans

$$\left.\begin{array}{rcl} a_1x^1 + \ldots + a_nx^n & = & b \\ \bar{a}_1x^1 + \ldots + \bar{a}_nx^n & = & \bar{b} \end{array}\right\}$$

són paral·lels si i només si tenen la mateixa direcció, és a dir, si les equacions $a_1x^1 + \ldots + a_nx^n = 0$, $\bar{a}_1x^1 + \ldots + \bar{a}_nx^n = 0$ tenen les mateixes solucions. Això

passa si els coeficients són proporcionals: $\exists c$ tal que $a_i = c\bar{a}_i$, $i = 1, \ldots, n$. Si, a més a més, $b = c\bar{b}$, els dos hiperplans coincideixen; en cas contrari, si $b \neq c\bar{b}$, no es tallen. (Compareu amb (4.2)).

II. Una recta i un hiperplà d'equacions

$$\frac{x^1 - a^1}{v^1} = \ldots = \frac{x^n - a^n}{v^n}, \qquad b_1 x^1 + \ldots + b_n x^n = b$$

són paral·lels si la direcció de la recta, $\langle (v^1, \ldots, v^n) \rangle$, està continguda en la direcció del pla; és a dir, si

$$b_1 v^1 + \ldots + b_n v^n = 0.$$

IX.10 Raó simple

Donats tres punts alineats a_1, a_2, a_3 d'un espai afí (A, E), es diu *raó simple de* a_1, a_2, a_3, i s'escriu $(a_1 a_2 a_3)$, l'element $r \in K$ tal que

$$\overrightarrow{a_1 a_3} = r \overrightarrow{a_1 a_2}.$$

$(a_1 a_2 a_3) = r$ està definit sempre que $a_1 \neq a_2$, és 0 si $a_1 = a_3$ i és 1 si $a_2 = a_3$.

Proposició 10.1 *Si $a_2 \neq a_3$ i (α, β) són les coordenades baricèntriques de a_1 en el sistema de referència de la recta $\{a_2, a_3\}$, llavors*

$$(a_1 a_2 a_3) = \frac{-\alpha}{\beta}.$$

DEMOSTRACIÓ: Per a cada p,

$$\overrightarrow{pa_1} = \alpha \overrightarrow{pa_2} + \beta \overrightarrow{pa_3} \text{ amb } \alpha + \beta = 1.$$

En particular, $0 = \alpha \overrightarrow{a_1 a_2} + \beta \overrightarrow{a_1 a_3}$, d'on $\overrightarrow{a_1 a_3} = -\frac{\alpha}{\beta} \overrightarrow{a_1 a_2}$. □

Corol·lari 10.2 *Suposem $K = \mathbf{R}$. a_1 és del segment $\overline{a_2 a_3}$ si i només si*

$$(a_1 a_2 a_3) < 0.$$

DEMOSTRACIÓ: $a_1 \in \overline{a_2 a_3} \Leftrightarrow \alpha, \beta > 0$ (§6) $\Leftrightarrow (a_1 a_2 a_3) < 0$. □

Exemples:

1. Considerem la recta afí (K, K) i siguin $x_1, x_2, x_3 \in K$ tres punts qualssevol. Recordem (§1) que $\overrightarrow{x_i x_j} = x_j - x_i \in K$. Per tant, si $x_1 \neq x_2$,

$$(x_1 x_2 x_3) = \frac{x_3 - x_1}{x_2 - x_1}.$$

2. En el cas particular $K = \mathbf{C}$, donats $z_1, z_2, z_3 \in \mathbf{C}$ amb $z_1 \neq z_2$,

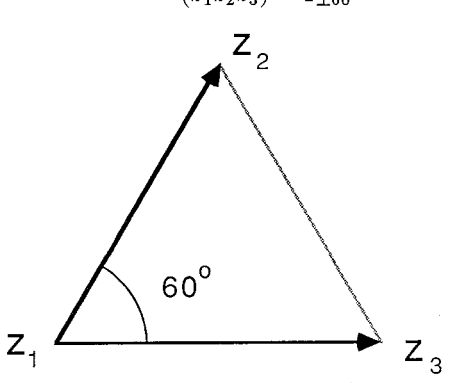
$$(z_1 z_2 z_3) = \frac{z_3 - z_1}{z_2 - z_1}.$$

Observem que el mòdul de la raó és el quocient dels mòduls de $z_3 - z_1$ i $z_2 - z_1$, i l'argument del complex $(z_1 z_2 z_3)$ és la diferència

$$\arg(z_3 - z_1) - \arg(z_2 - z_1).$$

Així, per exemple, z_1, z_2, z_3 formen un triangle equilàter si i només si

$$(z_1 z_2 z_3) = 1_{\pm 60°}$$



Teorema 10.3 (de Menelao) *Siguin a_0, a_1, a_2 punts linealment independents d'un espai afí (A, E) i siguin b_0, b_1, b_2 punts de les rectes determinades per $a_1 a_2$, $a_0 a_2$ i $a_0 a_1$ respectivament, alineats i diferents de a_0, a_1, a_2. Aleshores*

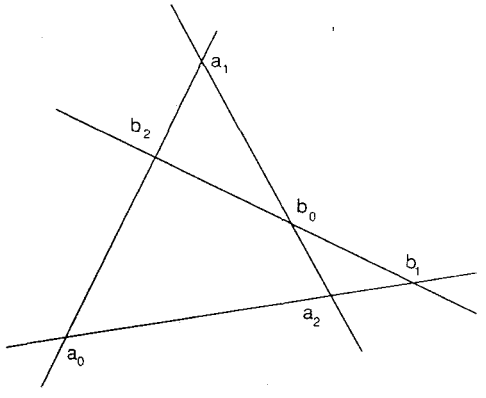
$$(b_0 a_1 a_2) \cdot (b_1 a_2 a_0) \cdot (b_2 a_0 a_1) = 1.$$

DEMOSTRACIÓ: Les coordenades baricèntriques dels punts b_0, b_1, b_2 en el sistema $\{a_0, a_1, a_2\}$ són del tipus

$$b_0 = (0, \alpha_0, \beta_0), \qquad b_1 = (\alpha_1, 0, \beta_1) \qquad b_2 = (\alpha_2, \beta_2, 0).$$

Aleshores (10.1) ens diu que

$$(b_0 a_1 a_2) = \frac{-\alpha_0}{\beta_0} \qquad (b_1 a_2 a_0) = \frac{-\beta_1}{\alpha_1} \qquad (b_2 a_0 a_1) = \frac{-\alpha_2}{\beta_2};$$



d'on

$$(b_0a_1a_2)\cdot(b_1a_2a_0)\cdot(b_2a_0a_1) = -\frac{\alpha_0\beta_1\alpha_2}{\beta_0\alpha_1\beta_2}.$$

Però, d'altra banda, vam veure al §7 que

$$b_0, b_1, b_2 \text{ alineats} \quad \Leftrightarrow \quad 0 = \begin{vmatrix} 0 & \alpha_1 & \alpha_2 \\ \alpha_0 & 0 & \beta_2 \\ \beta_0 & \beta_1 & 0 \end{vmatrix} = \alpha_0\beta_1\alpha_2 + \alpha_1\beta_2\beta_0$$

$$\Leftrightarrow \quad -\frac{\alpha_0\beta_1\alpha_2}{\alpha_1\beta_2\beta_0} = 1.$$

Això demostra el teorema. □

Corol·lari 10.4 *Suposem $K = \mathbf{R}$. Amb les notacions del teorema anterior, no és possible*

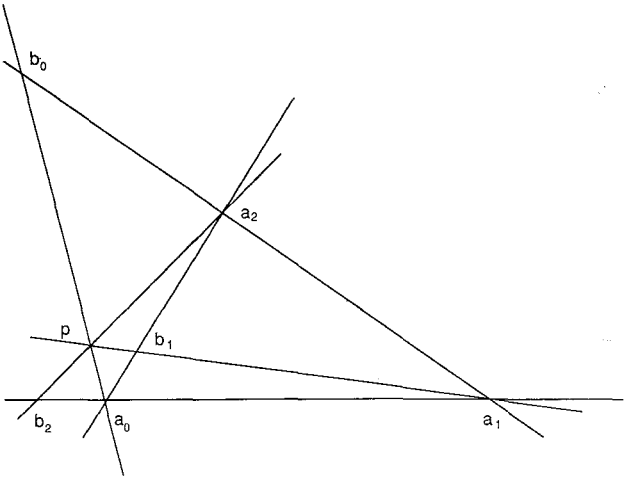
$$b_0 \in \overline{a_1a_2}, \qquad b_1 \in \overline{a_2a_0}, \qquad b_3 \in \overline{a_0a_1}$$

simultàniament. En altres paraules, una recta real no pot tallar simultàniament els tres "costats" d'un triangle.

DEMOSTRACIÓ: És conseqüència de 10.3 i 10.2. □

Teorema 10.5 (de Ceva) *Siguin a_0, a_1, a_2 punts linealment independents d'un espai afí de dimensió 2 (A, E). Donat $p \in A$ diferent de a_0, a_1, a_2 , designem per b_i $(i = 0, 1, 2)$ la intersecció de la recta determinada per pa_i amb la recta determinada pels altres dos punts a_ka_h $(k, h \neq i)$. Aleshores*

$$(b_0a_1a_2)\cdot(b_1a_2a_0)\cdot(b_2a_0a_1) = -1.$$



DEMOSTRACIÓ: Si (α, β, γ) són les coordenades baricèntriques de p en el sistema $\{a_0, a_1, a_2\}$, resulta que

$$\begin{aligned} &b_0 \text{ té per coordenades } (0, \tfrac{\beta}{1-\alpha}, \tfrac{\gamma}{1-\alpha}), \text{ d'on } (b_0 a_1 a_2) = -\tfrac{\beta}{\gamma}; \\ &b_1 \text{ té per coordenades } (\tfrac{\alpha}{1-\beta}, 0, \tfrac{\gamma}{1-\beta}), \text{ d'on } (b_1 a_2 a_0) = -\tfrac{\gamma}{\alpha}; \\ &b_2 \text{ té per coordenades } (\tfrac{\alpha}{1-\gamma}, \tfrac{\beta}{1-\gamma}, 0), \text{ d'on } (b_2 a_0 a_1) = -\tfrac{\alpha}{\beta}. \end{aligned}$$

Per tant,

$$(b_0 a_1 a_2) \cdot (b_1 a_2 a_0) \cdot (b_2 a_0 a_1) = -\frac{\beta\gamma\alpha}{\gamma\alpha\beta} = -1. \square$$

IX.11 Orientació d'un espai afí real

Sigui E un espai vectorial de dimensió n sobre $\mathbf{R}$. Direm que dues bases $\{e_1, \ldots, e_n\}$ i $\{u_1, \ldots, u_n\}$ són de la *mateixa orientació* si

$$\det_{(e_i)}(u_1, \ldots, u_n) > 0.$$

En cas contrari direm que són *d'orientacions oposades.* "Tenir la mateixa orientació" és un relació d'equivalència en el conjunt de bases de E, i dóna lloc a dues classes d'equivalència que denominarem *orientacions* de E. *Orientar l'espai vectorial* E és escollir una d'aquestes dues orientacions. L'orientació escollida es diu *orientació positiva* i l'altra *orientació negativa.*

Exemple:

Si escollim com a orientació positiva la de e_1, e_2, e_3, llavors e_2, e_1, e_3 és d'orientació negativa però e_2, e_3, e_1 és d'orientació positiva.

Un automorfisme $f : E \longrightarrow E$ *conserva l'orientació* si les bases $e_1, \ldots, e_n$ i $f(e_1), \ldots, f(e_n)$ són de la mateixa orientació: és a dir, si

$$\det_{(e_i)}(f(e_1), \ldots, f(e_n)) = \det f > 0.$$

La definició anterior és, doncs, independent de la base $e_1, \ldots, e_n$.

Si $\det f < 0$ direm que f *inverteix l'orientació.*

Orientar un espai afí (A, E) és escollir una orientació de E. *Orientar una varietat lineal* $a + F$ és escollir una orientació de F.

IX.12 Semiespais

En un espai afí real (A, E) de dimensió n cada hiperplà H divideix la resta de punts de A en dues zones de la següent manera: considerem a $A - H$ la relació

$$p \sim q \Leftrightarrow \text{ el segment } \overline{pq} \text{ no talla } H.$$

Les propietats reflexiva ($p \sim p$, $\forall p \in A - H$) i simètrica ($p \sim q \Rightarrow q \sim p$) d'aquesta relació són òbvies. Abans de demostrar la propietat transitiva anem a caracteritzar la relació $p \sim q$ d'una altra forma. Sigui $a_1 x^1 + \ldots + a_n x^n + b = 0$ l'equació de l'hiperplà H en un cert sistema de referència cartesià. Per a tot punt $y \in A - H$ de coordenades $(y^1, \ldots, y^n)$ designarem per a_y el nombre real

$$a_y = a_1 y^1 + \ldots + a_n y^n + b \neq 0.$$

Si el segment

$$\overline{pq} = \{x \in A, x^i = (1-t)p^i + tq^i, i = 1, \ldots, n, 0 \leq t \leq 1\}$$

talla H, existeix un valor de t, $0 < t < 1$, tal que

$$a_1\left((1-t)p^1 + tq^1\right) + \ldots + a_n\left((1-t)p^n + tq^n\right) + b = 0.$$

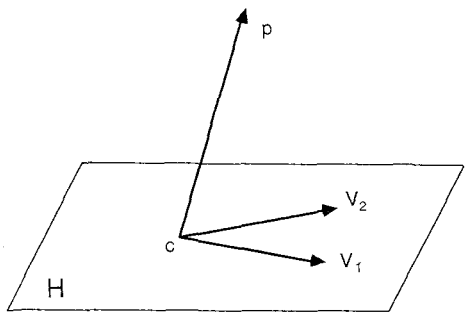
Podem escriure aquesta expressió així

$$(1-t)a_p + ta_q = 0,$$

d'on $0 < t = \dfrac{a_p}{a_p - a_q} < 1$, i això és cert si i només si a_p i a_q són un positiu i l'altre negatiu. És a dir,

$$p \sim q \Leftrightarrow a_p \cdot a_q > 0.$$

D'aquí resulta immediatament que aquesta relació és d'equivalència i que divideix $A - H$ en dos subconjunts que anomenarem *semiespais* de A.



La noció de semiespai està relacionada amb el concepte d'orientació definit al §11. En efecte, sigui $v_1, \ldots, v_{n-1}$ una base de la direcció de H i sigui $c \in H$. Posem

$$a_x = \begin{vmatrix} x^1 - c^1 & v_1^1 & \ldots & v_{n-1}^1 \\ \vdots & \vdots & & \vdots \\ x^n - c^n & v_1^n & \ldots & v_{n-1}^n \end{vmatrix} = \det(\overrightarrow{cx}, v_1, \ldots, v_{n-1}).$$

Així $a_x = 0$ és una equació de H i, si $p \in A - H$,

$$a_p = \det(\overrightarrow{cp}, v_1, \ldots, v_{n-1}) \neq 0.$$

Dos punts p, q són, per tant, del mateix semiespai si i només si les bases

$$\{\overrightarrow{cp}, v_1, \ldots, v_{n-1}\} \text{ i } \{\overrightarrow{cq}, v_1, \ldots, v_{n-1}\}$$

són de la mateixa orientació.

IX.13 Nota històrica

Pierre de Fermat (1601-1665) i René Descartes (1596-1650), obsessionats per la necessitat de mètodes quantitatius a la geometria i impressionats pel poder de l'àlgebra, van iniciar l'aplicabilitat d'aquesta en l'estudi de la geometria, creant els sistemes de coordenades associant equacions algebraiques a corbes i superfícies. Aquesta idea ha estat una de les més riques i fructíferes en el desenvolupament de les matemàtiques.

Tant Fermat com Descartes estaven motivats per les necessitats de la ciència i per un interès en la metodologia. Especialment Descartes (el primer gran filòsof modern, un fundador de la biologia moderna, un físic de categoria i matemàtic només incidentalment) va fer de la metodologia l'objectiu principal de tota la seva obra.

Les primeres nocions nebuloses d'un hiperespai de dimensió $n > 3$ es perden en l'obscuritat del passat i es barregen amb consideracions metafísiques. El primer article científic que tracta explícitament el tema es deu a Arthur Cayley (1821-1895) i és de 1843. El segueixen una sèrie de treballs d'ell mateix, de James Joseph Sylvester (1814-1897) i de William Kingdom Clifford (1845-1879) a Anglaterra i de Hermann Günther Grassmann (1809-1877) i Ludwig Schläfli (1814-1895) al continent. La introducció de coordenades es fa durant la segona meitat del segle 19 via l'estudi dels espais aritmètics.

El 1818 August Ferdinand Möbius (1790-1868) ja havia tingut la idea d'una anàlisi geomètrica en els espais de dimensió 2 i 3, que desenvolupà a partir del 1823 sota el nom de "càlcul baricèntric", inspirat per la teoria de centres de gravetat. És el que avui anomenem un sistema de coordenades baricèntriques. No és però fins a finals de segle que Schläfli i Camille Jordan (1838-1922) desenvolupen explícitament les nocions de la geometria afí (i de l'euclidiana) de dimensió n. La linealització de la geometria és un fet.

IX.14 Exercicis

1. En plena guerra, el servei de contraespionatge dels bons intercepta un missatge dels dolents que diu: "El dia 23 del proper mes de febrer, a les 6.25 hores p.m. iniciarem un atac contra els bons. Començarà amb el foc intens d'una bateria d'artilleria de campanya que tenim situada al punt $(2.3, -5)$. Al punt $(-11, 0.3)$ tenim preparada un divisió d'infanteria disposada a entrar en combat després del bombardeig inicial i, finalment, al punt $(9, 7.1)$ tenim una bateria d'artilleria antiaèria per defensar les unitats suara citades de l'atac dels avions bons. Dirigirà l'operació des de l'origen de coordenades el general Bum-Bum."

El cap de defensa dels bons ordena urgentment que surtin escamots de reconeixement a localitzar les unitats enemigues. Les patrulles dels bons descobreixen que la bateria d'artilleria de campanya està situada al punt $(15, 7.5)$, la divisió d'infanteria està acampada al punt $(19, -2.7)$ i la bateria d'artilleria antiaèria s'ha aposentat al punt $(-14, 9.2)$. Totes les dades segons el sistema de referència dels bons, és clar. Un cop rebuda aquesta informació el general bo fa que li portin un matemàtic i li pregunta si és possible esbrinar quin és l'origen de coordenades dels dolents, per tal de carregar-se el general Bum-Bum i desmoralitzar així els dolents, els quals molt probablement ja no durien a terme l'ofensiva. El matemàtic es recordava que quan feia primer havia resolt un problema semblant a la classe d'Àlgebra i Geometria i en un no res li va calcular el tan desitjat punt. Com s'ho va fer ?

2. a) Demostreu que a $\mathbf{R}^2$ els punts mitjos de qualsevol quadrilàter formen un paral·lelogram.

b) Demostreu que un quadrilàter és un paral·lelogram si i només si les diagonals es tallen al punt mig.

3. Siguin $A_i = (a_i, b_i, c_i)$, $i = 1, 2, 3, 4$, punts d'un espai afí de dimensió 3. Demostreu que els A_i són coplanaris si i només si

$$\begin{vmatrix} a_1 & b_1 & c_1 & 1 \\ a_2 & b_2 & c_2 & 1 \\ a_3 & b_3 & c_3 & 1 \\ a_4 & b_4 & c_4 & 1 \end{vmatrix} = 0.$$

4. Siguin $a_1, \ldots, a_n$ punts d'un espai afí. Demostreu que les rectes que uneixen cada a_i, $i = 1, \ldots, n$, amb el baricentre dels altres punts són concurrents (en el baricentre de $a_1, \ldots, a_n$).

5. Donat un triangle abc del pla afí real i tres punts a', b', c' sobre els costats bc, ca, ab respectivament, trobeu una condició necessària i suficient perquè els triangles a, b, c i a', b', c' tinguin el mateix baricentre.

6. Calculeu les sis raons simples que s'obtenen en permutar tres punts alineats.

7. Sigui A una recta afí sobre $\mathbf{R}$ o $\mathbf{C}$. En quines condicions les sis raons simples de tres punts de A prenen com a màxim tres valors diferents?

8. (Teorema de Desargues). Considerem dos triangles ABC i $A'B'C'$ del pla afí real i denotem per a, b, c (resp. a', b', c') les rectes BC, AC, AB (resp. $B'C'$, $A'C'$, $A'B'$). Demostreu que les rectes AA', BB' i CC' són paral·leles o concurrents si i només si els punts $a \cap a'$, $b \cap b'$ i $c \cap c'$ estan alineats.

9. (Teorema de Papus). Siguin A, B, C i A', B', C' duesternes de punts alineats i denotem per a, b, c (resp. a', b', c') les rectes BC', AC', AB' (resp. $B'C$, $A'C$, $A'B$). Demostreu que els punts $a \cap a'$, $b \cap b'$ i $c \cap c'$ estan alineats.

10. Demostreu que un subconjunt d'un espai afí és una varietat lineal si i només si conté, amb cada parella de punts, la recta que determinen.

11. En el pla afí $A = \mathbf{Z}/(3) \times \mathbf{Z}/(3)$ sobre el cos $\mathbf{Z}/(3)$,

a) Quants punts hi ha ?

b) Quantes rectes hi ha ?

c) Quants punts té cada recta ?

d) Quantes rectes hi ha paral·leles a una donada ?

e) Quants feixos diferents de rectes paral·leles hi ha ?

12. Sigui A un pla afí sobre el cos $\mathbf{Z}/(5)$. Estudieu la posició relativa de les rectes

$$r : 3x + 2y + 2 = 0,\ s : 2x + y = 0,\ t : x - 3y + 1 = 0.$$

13. Siguin $A_i x + b_i = \vec{0}$, $i = 1, 2$, les equacions de dues varietats lineals. Demostreu que les equacions de la varietat intersecció s'obtenen reduint pel mètode de Gauss la matriu

$$\begin{pmatrix} A_1 & b_1 \\ A_2 & b_2 \end{pmatrix}.$$

Si el sistema és incompatible les dues varietats no es tallen i

a) si rang $\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ = màx (rang A_1, rang A_2), les varietats són paral·leles,

b) si rang $\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ > rang A_i, $i = 1, 2$, les varietats es creuen.

14. Siguin $Ax + b = \vec{0}$ amb $A \in M_{(n-k)\times n}(K)$, $b \in K^{n-k}$, rang $A = n - k$, les equacions d'una varietat lineal de dimensió k en un espai afí de dimensió n, en un sistema de referència cartesià.

Si $\bar{A}\bar{x} + \bar{b} = \vec{0}$ són les equacions en un altre sistema de referència, demostreu que, aleshores,

$$\begin{cases} \bar{A} = AV \\ \bar{b} = b + Aq \end{cases}$$

on V és la matriu del canvi de base i q les coordenades de l'origen.

15. Donades tres rectes del pla afí real d'equacions

$$3x + 2y = 1,\ y = 5,\ 6x + y = -13,$$

trobeu els triangles abc que tenen les seves mitjanes (XIII.11) sobre aquestes rectes, el vèrtex a sobre la primera recta i el punt $(-1, 2)$ com a punt mig del costat bc.

16. Les cares d'un tetràedre $abcd$ són tallades per una recta en quatre punts a', b', c', d'. Proveu que els punts mitjos dels segments aa', bb', cc', dd' són coplanaris.

17. Pels vèrtexs d'un tetràedre $abcd$ de l'espai afí real es tracen quatre rectes paral·leles que tallen les cares oposades en els punts a', b', c', d'. Determineu $k \in \mathbf{R}$ de manera que els punts que pertanyen als segments aa', bb', cc', dd' amb raó k siguin coplanaris.

18. A l'espai afí real es consideren tres rectes que es creuen dos a dos i són paral·leles a un pla. Demostreu que tota recta que talli les tres és paral·lela a un pla fix. Determineu aquest pla.

IX.15 Exercicis de programar

19. Siguin $p_0, \ldots, p_n$ punts donats de l'espai afí real $\mathbf{R}^n$. Feu un programa que permeti:

a) Comprovar que són linealment independents.

b) Donat un punt $x \in \mathbf{R}^n$ qualsevol, trobar les seves coordenades baricèntriques respecte al sistema $\{p_0, \ldots, p_n\}$.

20. Com a aplicació de l'exercici 19, prepareu tres programes que permetin:

a) Decidir si un punt x donat del pla afí real és interior, exterior o està en un costat del triangle determinat per tres punts p_0, p_1, p_2. (Veure §6, exemple 3).

b) El mateix a l'espai afí amb quatre punts linealment independents.

c) Donat un polígon convex del pla afí per la successió ordenada dels seus vèrtexs, decidir si un punt x donat és interior, exterior o està en un costat. (Indicació: descomponeu el polígon en triangles amb un vèrtex comú).

21. Sigui $\{p; e_1, \ldots, e_n\}$ el sistema de referència cartesià canònic de l'espai afí $\mathbf{R}^n$ ($p = (0, \ldots, 0)$, $e_i = (0, \ldots, 1, \ldots, 0)$, $i = 1, \ldots, n$). Sigui $\{q; v_1, \ldots, v_n\}$ un altre sistema de referència donat.

a) Feu un programa que transporti les coordenades d'un punt donat d'un sistema a l'altre. (Caldrà utilitzar l'exercici VII.12 per invertir la matriu $\begin{pmatrix} V & q \\ 0 & 1 \end{pmatrix}$, on V és la matriu del canvi de base i q les coordenades del nou origen).

b) Feu un programa que canviï de sistema de referència les equacions cartesianes d'una varietat lineal. (§9 i exercici 14).

22. Donades les equacions cartesianes de dues varietats lineals de $\mathbf{R}^n$ en el sistema de referència canònic, feu un programa que permeti decidir si es tallen, es creuen o són paral·leles. Si es tallen doneu les equacions simplificades de la intersecció. (Veure exercici 13).

23. Feu un programa que calculi la raó simple de tres punts alineats de $\mathbf{R}^n$ i estudiï l'efecte de permutar els punts.

Capítol X

Afinitats

De la mateixa manera que en el capítol de grups consideràvem les aplicacions entre ells que "respectaven" llurs operacions (els homomorfismes), i que en estudiar espais vectorials ens interessàvem per les aplicacions que conservaven les dues operacions (les aplicacions lineals), volem ara estudiar aplicacions entre espais afins que respectin llur estructura afí: les aplicacions afins o afinitats.

X.1 Definició i primeres propietats

Siguin (A_1, E_1, φ_1) i (A_2, E_2, φ_2) dos espais afins sobre el mateix cos K. Una *aplicació afí* o *afinitat* entre aquests espais és una aplicació

$$f : A_1 \longrightarrow A_2$$

i una aplicació lineal

$$\tilde{f} : E_1 \longrightarrow E_2$$

tals que el diagrama d'aplicacions

$$\begin{array}{ccc} A_1 \times A_1 & \xrightarrow{\varphi_1} & E_1 \\ {\scriptstyle f\times f}\downarrow & & \downarrow \tilde{f} \\ A_2 \times A_2 & \xrightarrow{\varphi_2} & E_2 \end{array}$$

commuta. Això vol dir que $\tilde{f} \circ \varphi_1 = \varphi_2 \circ (f \times f)$. És a dir,

$$\forall a, b \in A_1, \qquad \tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)}.$$

Recordem que l'estructura d'un espai afí pot ésser donada també pel conjunt de les seves translacions: $T_1 : A_1 \times E_1 \longrightarrow A_1$ i $T_2 : A_2 \times E_2 \longrightarrow A_2$. Aleshores

seria lògic exigir que les aplicacions f i $\tilde{f}$ fessin commutatiu el diagrama

$$\begin{array}{ccc} A_1 \times E_1 & \xrightarrow{T_1} & A_1 \\ f\times\tilde{f}\downarrow & & \downarrow f \\ A_2 \times E_2 & \xrightarrow{T_2} & A_2 \ ; \end{array}$$

és a dir, $f \circ T_1 = T_2 \circ (f \times \tilde{f})$. Això equival a dir que $\forall a \in A$ i $\forall u \in E_1$

$$f(a+u) = f(a) + \tilde{f}(u).$$

La proposició següent demostra que aquesta condició equival a la que hem imposat en la definició.

Proposició 1.1 *Les condicions*

i) $\tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)} \qquad \forall a, b \in A_1$

ii) $f(a+u) = f(a) + \tilde{f}(u) \qquad \forall a \in A_1, \forall u \in E_1$

són equivalents.

DEMOSTRACIÓ: i) $\Rightarrow$ ii). En efecte, donats $a \in A_1$, $u \in E_1$, sigui $b \in A_1$ tal que $\overrightarrow{ab} = u$. Llavors, per i), $\tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)}$ i, per tant,

$$f(a+u) = f(b) = f(a) + \tilde{f}(\overrightarrow{ab}) = f(a) + \tilde{f}(u).$$

ii) $\Rightarrow$ i). En efecte, donats $a, b \in A_1$, de $b = a + \overrightarrow{ab}$ es dedueix $f(b) = f(a) + \tilde{f}(\overrightarrow{ab})$ i, per tant,

$$\overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}). \square$$

Per abús de llenguatge es diu que una aplicació $f : A_1 \longrightarrow A_2$ és una afinitat si existeix una aplicació lineal $\tilde{f} : E_1 \longrightarrow E_2$ que compleix una de les condicions de (1.1). $\tilde{f}$ es diu l'*aplicació lineal associada* a f.

Anem a demostrar ara uns quants fets que es dedueixen immediatament de les definicions.

Proposició 1.2 *Siguin $f, g : A_1 \longrightarrow A_2$ dues afinitats que coincideixen sobre un punt p, $f(p) = g(p)$, i que tenen la mateixa aplicació lineal associada, $\tilde{f} = \tilde{g}$. Aleshores $f = g$.*

DEMOSTRACIÓ: Per a tot $a \in A_1$,

$$f(a) = f(p + \overrightarrow{pa}) = f(p) + \tilde{f}(\overrightarrow{pa}) = g(p) + \tilde{g}(\overrightarrow{pa}) = g(p + \overrightarrow{pa}) = g(a). \square$$

Proposició 1.3 *Donada una aplicació lineal $\varphi : E_1 \longrightarrow E_2$ i un parell de punts $p \in A_1, q \in A_2$, existeix una afinitat f, i només una, tal que $f(p) = q$ i $\tilde{f} = \varphi$.*

DEMOSTRACIÓ: Si existeix, f és única per (1.2). Definim $f : A_1 \longrightarrow A_2$ per

$$f(a) = q + \varphi(\overrightarrow{pa}), \qquad \forall a \in A_1.$$

En particular $f(p) = q$ i $\overrightarrow{f(a)f(b)} = \overrightarrow{qf(b)} - \overrightarrow{qf(a)} = \varphi(\overrightarrow{pb}) - \varphi(\overrightarrow{pa}) = \varphi(\overrightarrow{ab}) \quad \forall a, b \in A_1$. Això demostra que f és una afinitat amb aplicació lineal associada φ que transforma p en q.□

Proposició 1.4 *Suposem A_1 de dimensió n. Donats $(n+1)$ punts linealment independents $a_0, \ldots, a_n$ de A_1, i $(n+1)$ punts arbitraris $b_0, \ldots, b_n$ de A_2, existeix una afinitat $f : A_1 \longrightarrow A_2$, i només una, tal que $f(a_i) = b_i$, $\quad i = 0, \ldots, n$.*

DEMOSTRACIÓ: Si existeix l'afinitat f, la seva aplicació lineal associada ha d'ésser l'única aplicació lineal $\tilde{f} : E_1 \longrightarrow E_2$ tal que $\tilde{f}(\overrightarrow{a_0a_i}) = \overrightarrow{b_0b_i}$ per a $i = 1, \ldots, n$. Definim doncs $f : A_1 \longrightarrow A_2$ com l'única aplicació afí que té per aplicació lineal associada aquesta $\tilde{f}$ i que transforma a_0 en b_0 : $f(a_0) = b_0$ (1.3). f compleix

$$f(a_i) = f(a_0 + \overrightarrow{a_0a_i}) = f(a_0) + \tilde{f}(\overrightarrow{a_0a_i}) = b_0 + \overrightarrow{b_0b_i} = b_i$$

i és, per tant, l'aplicació buscada.□

Proposició 1.5 *Si $f : A_1 \longrightarrow A_2$ i $g : A_2 \longrightarrow A_3$ són afinitats, aleshores $g \circ f : A_1 \longrightarrow A_3$ és una afinitat, i $\widetilde{g \circ f} = \tilde{g} \circ \tilde{f}$.*

DEMOSTRACIÓ: $\forall a, b \in A_1$,

$$\tilde{g} \circ \tilde{f}(\overrightarrow{ab}) = \tilde{g}(\overrightarrow{f(a)f(b)}) = \overrightarrow{g \circ f(a)\ g \circ f(b)}.$$

L'aplicació $\tilde{g} \circ \tilde{f}$ compleix, per tant, $i)$ de (1.1) i és l'aplicació lineal associada a $g \circ f$.□

La proposició següent demostra que hi ha una relació molt estreta entre una afinitat i la seva aplicació lineal associada.

Proposició 1.6 *Donada una afinitat $(f, \tilde{f})$ de (A_1, E_1) en (A_2, E_2),*
a) f és injectiva $\Leftrightarrow$ $\tilde{f}$ és injectiva;
b) f és exhaustiva $\Leftrightarrow$ $\tilde{f}$ és exhaustiva;
c) f és bijectiva $\Leftrightarrow$ $\tilde{f}$ és bijectiva.

DEMOSTRACIÓ: Suposem f injectiva. Donat $v \in \operatorname{Nuc} \tilde{f}$, si $v = \overrightarrow{ab}$, tenim

$$\vec{0} = \tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)} \quad \Rightarrow \quad f(a) = f(b) \quad \Rightarrow \quad a = b \quad \Rightarrow \quad v = \overrightarrow{ab} = \vec{0}.$$

Per tant, $\operatorname{Nuc} \tilde{f} = \{\vec{0}\}$ i $\tilde{f}$ és injectiva.

Suposem $\tilde{f}$ injectiva. Llavors,

$$f(a) = f(b) \Rightarrow \vec{0} = \overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}) \Rightarrow$$
$$\Rightarrow \overrightarrow{ab} \in \operatorname{Nuc} \tilde{f} = \{\vec{0}\} \Rightarrow \overrightarrow{ab} = \vec{0} \Rightarrow a = b.$$

Tenim així demostrat a).

Suposem f exhaustiva. Donat $u \in E_2$, posem $u = \overrightarrow{cd}$. Siguin $a, b \in A_1$, tals que $f(a) = c, f(b) = d$. Llavors $u = \overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab})$; per tant, $\tilde{f}$ és exhaustiva.

Suposem $\tilde{f}$ exhaustiva. Donat $c \in A_2$, considerem un vector $u = \overrightarrow{f(a)c}$, on a és un punt qualsevol de A_1. Per ésser $\tilde{f}$ exhaustiva existeix $v \in E_1$ tal que $\tilde{f}(v) = u$. Escrivim $v = \overrightarrow{ab}$. Llavors $\overrightarrow{f(a)c} = u = \tilde{f}(v) = \overrightarrow{f(a)f(b)}$, d'on $f(b) = c$. Per tant, f també és exhaustiva. Això demostra b).

L'afirmació c) és conseqüència de les dues anteriors. □

Corol·lari 1.7 *Si f és una afinitat bijectiva amb aplicació lineal associada $\tilde{f}$, llavors f^{-1} és una afinitat amb aplicació lineal associada $\tilde{f}^{-1}$.*

DEMOSTRACIÓ: En virtut de la proposició (1.6) només resta demostrar que $\tilde{f}^{-1}(\overrightarrow{cd}) = \overrightarrow{f^{-1}(c)f^{-1}(d)}$; però això es dedueix immediatament de

$$\tilde{f}(\overrightarrow{f^{-1}(c)f^{-1}(d)}) = \overrightarrow{ff^{-1}(c)ff^{-1}(d)} = \overrightarrow{cd}. \square$$

D'una afinitat bijectiva en direm *isomorfisme afí*. Dos espais afins són *isomorfs* si existeix un isomorfisme afí entre ells. De (1.6) resulta que, si A_1 i A_2 són isomorfs, els seus espais vectorials associats E_1, E_2 també ho són. Recíprocament, si $\varphi : E_1 \longrightarrow E_2$ és un isomorfisme, (1.3) i (1.6) aseguren l'existència d'un isomorfisme afí $f : A_1 \longrightarrow A_2$. Si els espais són de dimensió finita obtenim, en particular, el següent resultat.

Proposició 1.8 *Dos espais afins de dimensió finita sobre el mateix cos són isomorfs si i només si tenen la mateixa dimensió.* □

De (1.8) es dedueix que tot espai afí de dimensió finita és isomorf a un espai afí estàndard K^n (IX.1).

X.2 Uns exemples

En aquest apartat estudiarem exemples d'aplicaciones afins d'un espai afí en ell mateix.

I Translacions

Sigui T_v una translació de vector $v \in E$. $\forall a, b \in A$.

$$\overrightarrow{T_v(a)T_v(b)} = \overrightarrow{T_v(a)a} + \overrightarrow{ab} + \overrightarrow{bT_v(b)} = -v + \overrightarrow{ab} + v = \overrightarrow{ab} = I_E(\overrightarrow{ab}),$$

on I_E indica l'aplicació identitat de E. Això demostra que T_v és una afinitat amb aplicació lineal associada I_E.

Recíprocament, si $f : A \longrightarrow A$ és una afinitat i $\tilde{f} = I_E$, $\forall a, b$

$$\overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}) = \overrightarrow{ab};$$

d'on

$$\overrightarrow{af(a)} = \overrightarrow{bf(b)}.$$

És a dir, el vector v determinat per un punt i la seva imatge és independent del punt escollit: $v = \overrightarrow{af(a)}$. Així doncs, per a tot a , $f(a) = a + v = T_v(a)$, d'on $f = T_v$.

Proposició 2.1 *Una afinitat $f : A \longrightarrow A$ és una translació si i només si $\tilde{f} = I_E$.*□

II Projeccions

Una afinitat $f : A \longrightarrow A$ es diu una *projecció* si $f^2 = f$.

Observem, primer de tot, que tot punt de $\operatorname{Im} f$ és fix: $ff(a) = f(a)$ $\forall a$. A més, si b és fix, $b = f(b) \in \operatorname{Im} f$. Així doncs, $\operatorname{Im} f$ és el conjunt de punts fixos de f.

Estudiem ara $\tilde{f}$. $\forall \overrightarrow{ab} \in E$

$$\tilde{f}^2(\overrightarrow{ab}) = \tilde{f}(\overrightarrow{f(a)f(b)}) = \overrightarrow{f^2(a)f^2(b)} = \overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}).$$

Així doncs, $\tilde{f}^2 = \tilde{f}$ i el polinomi mínim de $\tilde{f}$ és un divisor de $x^2 - x$. Apliquem els mètodes de (VIII.5) a l'estudi de $\tilde{f}$. Es poden donar tres casos:

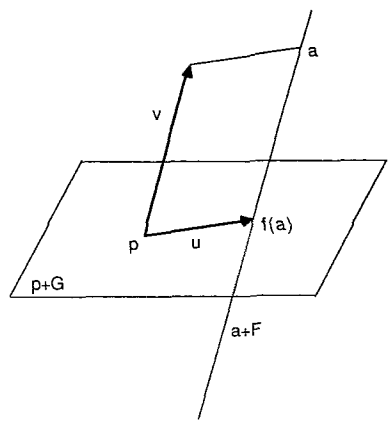
- Si el polinomi mínim de $\tilde{f}$ és x, $\tilde{f} = 0$ i $\overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}) = \vec{0}$ $\forall a, b \in A$. Així doncs, $f(a) = f(b)$ $\forall a, b \in A$ i, per tant, f aplica tot A en el mateix punt.

- Si el polinomi mínim de $\tilde{f}$ és $x - 1$, $\tilde{f} = I_E$ i f és una translació (2.1). Ara bé, f té punts fixos; per tant, és una translació de vector $\vec{0}$; és a dir, f és la identitat de A.

- Si el polinomi mínim de $\tilde{f}$ és $x(x-1)$, $E = F \oplus G$, on F és un subespai invariant sobre el qual $\tilde{f} = 0$ i G és un subespai invariant sobre el qual $\tilde{f} = I_G$ (VIII.4.4). Aleshores, si p és un punt fix,

$$f(a) = f(p + \overrightarrow{pa}) = p + \tilde{f}(\overrightarrow{pa}) \qquad \forall a \in A.$$

Si $\overrightarrow{pa} = v + u$ amb $v \in F$ i $u \in G$, $f(a) = p + u$. El conjunt de punts fixos de f és en aquest cas $p + G$ (3.3). Per altra banda, $\overrightarrow{f(a)a} = \overrightarrow{f(a)p} + \overrightarrow{pa} = -u + (v + u) = v$, d'on $f(a) \in a + F$. El punt $f(a)$ és, per tant, la intersecció de $a + F$ i $\operatorname{Im} f = p + G$. Observem que $E = F \oplus G$ implica que aquesta intersecció es redueix sempre a un sol punt (IX.4.4).



III Simetries

Una afinitat $f : A \longrightarrow A$ es diu una *simetria* si $f^2 = I_A$.

Suposem que el cos K no és de característica 2 (IX.6); els punts mitjos dels parells formats per un punt a i la seva imatge $f(a)$ són fixos:

$$m = \tfrac{1}{2}a + \tfrac{1}{2}f(a) \quad \Leftrightarrow \quad \overrightarrow{am} = \tfrac{1}{2}\overrightarrow{af(a)} \quad \Leftrightarrow \quad \overrightarrow{f(a)f(m)} = \tfrac{1}{2}\overrightarrow{f(a)a}$$
$$\Leftrightarrow \quad f(m) = \tfrac{1}{2}f(a) + \tfrac{1}{2}a = m.$$

Estudiem ara $\tilde{f}$. Com que $f^2 = I_A$, també $\tilde{f}^2 = I_E$ i el polinomi mínim de $\tilde{f}$ és un divisor de $x^2 - 1$ (VIII.4. Exemple 2). Hi ha tres possibilitats:

- Si el polinomi mínim de $\tilde{f}$ és $(x - 1)$, $\tilde{f} = I_E$ i f és una translació amb punts fixos; per tant, $f = I_A$.
- Si el polinomi mínim de $\tilde{f}$ és $(x + 1)$, $\tilde{f} = -I_E$. Si p és un punt fix, aleshores $\forall a \quad \overrightarrow{pf(a)} = \tilde{f}(\overrightarrow{pa}) = -\overrightarrow{pa}$, i això equival a $p = \frac{1}{2}a + \frac{1}{2}f(a)$. Existeix doncs un únic punt fix p que és punt mig del parell $a, f(a)$ $\forall a$. f es diu llavors una *simetria central* de *centre* p.
- Si el polinomi mínim de $\tilde{f}$ és $(x - 1)(x + 1)$, $E = F \oplus G$, on F és un subespai invariant sobre el qual $\tilde{f}$ és I_F, i G és un subespai invariant sobre el qual $\tilde{f}$ és

$-I_G$. Sigui p un punt fix; $\forall a \in A$, si $\overrightarrow{pa} = u + v$ amb $u \in F$ i $v \in G$,

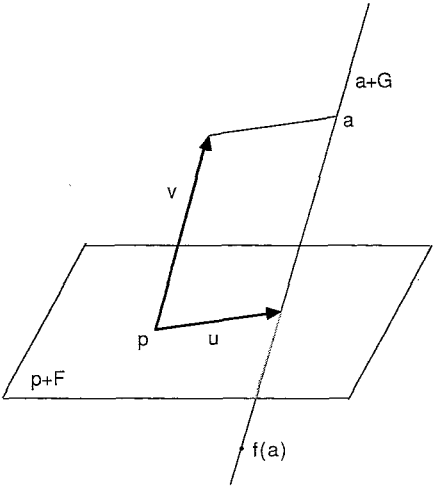
$$f(a) = f(p + \overrightarrow{pa}) = p + \tilde{f}(u + v) = p + u - v.$$

D'aquí resulta que els punts fixos de f són de la varietat $p + F$ (Prop. 3.3). Per altra banda,

$$\overrightarrow{f(a)a} = 2v \Rightarrow f(a) \in a + G$$

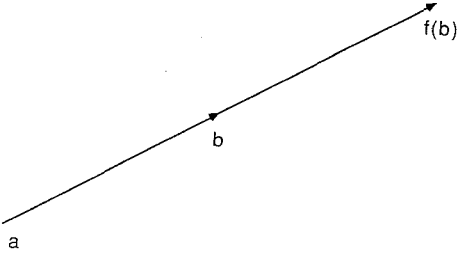
$$u = \frac{1}{2}\overrightarrow{pa} + \frac{1}{2}\overrightarrow{pf(a)} \quad \Rightarrow \quad \frac{1}{2}a + \frac{1}{2}f(a) = p + u \in p + F,$$

i aquestes dues condicions determinen f(a).



IV Homotècies

Una afinitat $f : A \longrightarrow A$ es diu una *homotècia de raó* $r \neq 0, 1$ si $\tilde{f} = rI_E$.



Per estudiar els possibles punts fixos, escollim un punt auxiliar $p \in A$. Aleshores $a \in A$ és fix si

$$a = f(a) = f(p + \vec{pa}) = f(p) + r\,\vec{pa} \quad \Leftrightarrow \quad \vec{pa} = \overrightarrow{pf(p)} + r\,\vec{pa} \quad \Leftrightarrow$$
$$\Leftrightarrow \quad \vec{pa} = \frac{1}{1-r}\overrightarrow{pf(p)} \quad \Leftrightarrow \quad a = p + \frac{1}{1-r}\overrightarrow{pf(p)}.$$

Aquest és, per tant, l'únic punt fix de f; és diu el *centre* de l'homotècia. La imatge de qualsevol altre punt $b \in A$ és

$$f(b) = a + r\,\vec{ab}.$$

X.3 Més propietats de les afinitats

Proposició 3.1 *Si $f : A_1 \longrightarrow A_2$ és una afinitat i $a + F$ és una varietat lineal de A_1,*

$$f(a + F) = f(a) + \tilde{f}(F).$$

Si $b + G$ és una varietat lineal de A_2 i $a \in f^{-1}(b + G)$,

$$f^{-1}(b + G) = a + \tilde{f}^{-1}(G).$$

DEMOSTRACIÓ: La primera part és una conseqüència immediata de la definició d'afinitat. Per demostrar la segona part observem que

$$f(a + \tilde{f}^{-1}(G)) = f(a) + G = b + G \quad \Rightarrow \quad f^{-1}(b + G) \supset a + \tilde{f}^{-1}(G).$$

Per altra banda,

$$c \in f^{-1}(b + G) \quad \Rightarrow \quad f(c) \in b + G = f(a) + G \quad \Rightarrow$$
$$\Rightarrow \quad \overrightarrow{f(a)f(c)} = \tilde{f}(\vec{ac}) \in G \quad \Rightarrow$$
$$\Rightarrow \quad \vec{cd} \in \tilde{f}^{-1}(G) \quad \Rightarrow \quad c \in a + \tilde{f}^{-1}(G),$$

d'on $f^{-1}(b + G) \subset a + \tilde{f}^{-1}(G)$. Això demostra la segona igualtat de l'enunciat.□

Corol·lari 3.2

a) $\operatorname{Im} f$ *és una varietat lineal de dimensió* $\operatorname{rang} \tilde{f}$.

b) f transforma varietats paral·leles en varietats paral·leles. En particular, f transforma punts alineats en punts alineats.□

Proposició 3.3 *Si el conjunt de punts fixos d'una afinitat $f : A \longrightarrow A$ no és buit, és una varietat lineal de direcció el subespai de vectors propis de valor propi 1 de $\tilde{f}$.*

DEMOSTRACIÓ: Sigui a un punt fix de f, $f(a) = a$, i E_1 el subespai de vectors propis de valor propi $+1$. $\forall u \in E_1$, $f(a+u) = f(a) + \tilde{f}(u) = a + u$; els punts de $a + E_1$ són, doncs, tots fixos. Recíprocament, si b és fix, $b = f(b) = f(a + \overrightarrow{ab}) = a + \tilde{f}(\overrightarrow{ab}) \Rightarrow \tilde{f}(\overrightarrow{ab}) = \overrightarrow{ab} \Rightarrow \overrightarrow{ab} \in E_1 \Rightarrow b \in a + E_1$.□

Proposició 3.4 *Si $f : A_1 \longrightarrow A_2$ és una afinitat i $x = \sum_{i=1}^{r} x^i a_i$ amb $x^1 + \cdots + x^r = 1$, aleshores $f(x) = \sum_{i=1}^{r} x^i f(a_i)$.*

DEMOSTRACIÓ: Sigui $p \in A_1$ un punt qualsevol. Tenim $\overrightarrow{px} = \sum_{i=1}^{r} x^i \overrightarrow{pa_i}$; d'on

$$\overrightarrow{f(p)f(x)} = \tilde{f}(\overrightarrow{px}) = \sum_{i=1}^{r} x^i \tilde{f}(\overrightarrow{pa_i}) = \sum_{i=1}^{r} x^i \overrightarrow{f(p)f(a_i)},$$

com volíem demostrar.□

(3.4) ens diu, en particular, que tota afinitat transforma el baricentre de r punts en el baricentre de llurs imatges.

Anem a provar ara que la propietat (3.4) és suficientment restrictiva per a caracteritzar les afinitats.

Proposició 3.5 *Una aplicació de conjunts $f : A_1 \longrightarrow A_2$ és una afinitat si i només si, sempre que $x^1 + \cdots + x^r = 1$,*

$$f(\sum_{i=1}^{r} x^i a_i) = \sum_{i=1}^{r} x^i f(a_i).$$

DEMOSTRACIÓ: El primer pas ha d'ésser definir el que serà l'aplicació lineal associada a f. Observem que la condició $\tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)}$ ens determina ja $\tilde{f}$. L'únic problema és que cada vector $u \in E_1$ admet moltes representacions de la forma $u = \overrightarrow{ab}$. Per evitar aquesta pluralitat fixem un punt $p \in A_1$ i prenguem tots els vectors amb origen p. Així doncs, definim

$$\begin{array}{rccc} \tilde{f}: & E_1 & \longrightarrow & E_2 \\ & u = \overrightarrow{px} & \longmapsto & \tilde{f}(u) = \overrightarrow{f(p)f(x)}. \end{array}$$

Provem que $\tilde{f}$ és lineal: donats $u = \overrightarrow{px}$, $v = \overrightarrow{py}$, sigui $u + v = \overrightarrow{pz}$; aleshores

$$\tilde{f}(u+v) = \tilde{f}(\overrightarrow{pz}) = \overrightarrow{f(p)f(z)}$$
$$\tilde{f}(u) + \tilde{f}(v) = \tilde{f}(\overrightarrow{px}) + \tilde{f}(\overrightarrow{py}) = \overrightarrow{f(p)f(x)} + \overrightarrow{f(p)f(y)}$$

i hauríem de veure que $\overrightarrow{f(p)f(z)} = \overrightarrow{f(p)f(x)} + \overrightarrow{f(p)f(y)}$. Això equival a veure que $f(z) = -f(p) + f(x) + f(y)$. La condició de l'enunciat ens diu que això serà cert si $z = -p + x + y$, és a dir, si $\overrightarrow{pz} = \overrightarrow{px} + \overrightarrow{py}$, que és precisament d'on hem partit.

Sigui ara $u = \overrightarrow{px}$, $ku = \overrightarrow{py}$. $\tilde{f}(u) = \overrightarrow{f(p)f(x)}$, $\tilde{f}(ku) = \overrightarrow{f(p)f(y)}$ i hauríem de veure que $\overrightarrow{f(p)f(y)} = k\,\overrightarrow{f(p)f(x)}$. Això vol dir que $f(y) = (1-k)f(p) + kf(x)$. N'hi ha prou amb veure que $y = (1-k)p + kx$, és a dir, que $\overrightarrow{py} = k\,\overrightarrow{px}$, i això és cert per hipòtesi.

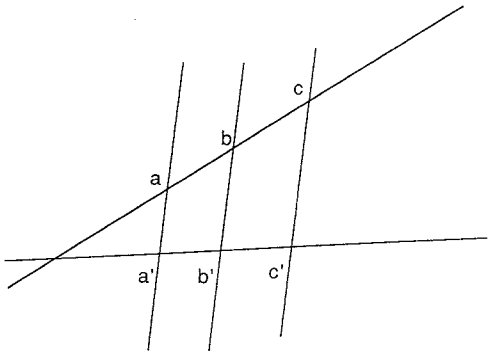
Només resta ara comprovar que $\tilde{f}$ és l'aplicació lineal associada a f. En efecte, $\forall a, b \in A$

$$\tilde{f}(\overrightarrow{ab}) = \tilde{f}(\overrightarrow{pb} - \overrightarrow{pa}) = \tilde{f}(\overrightarrow{pb}) - \tilde{f}(\overrightarrow{pa}) = \overrightarrow{f(p)f(b)} - \overrightarrow{f(p)f(a)} = \overrightarrow{f(a)f(b)}.\square$$

Proposició 3.6 *Les afinitats conserven la raó simple.*

DEMOSTRACIÓ: Sigui $r = (a_1a_2a_3)$. Tenim

$$\overrightarrow{a_1a_3} = r\overrightarrow{a_1a_2} \quad \Rightarrow \quad \overrightarrow{f(a_1)f(a_3)} = r\overrightarrow{f(a_1)f(a_2)} \quad \Leftrightarrow \quad (f(a_1)f(a_2)f(a_3)) = r.\square$$



Aquesta proposició, en el cas particular en què f sigui una projecció (§2,II), es coneix com el *teorema de Tales.*

La propietat (3.6) també caracteritza les afinitats.

Proposició 3.7 *Si el cos K no és de característica 2, una aplicació $f : A_1 \longrightarrow A_2$ és una afinitat si i només si conserva punts alineats i llurs raons simples.*

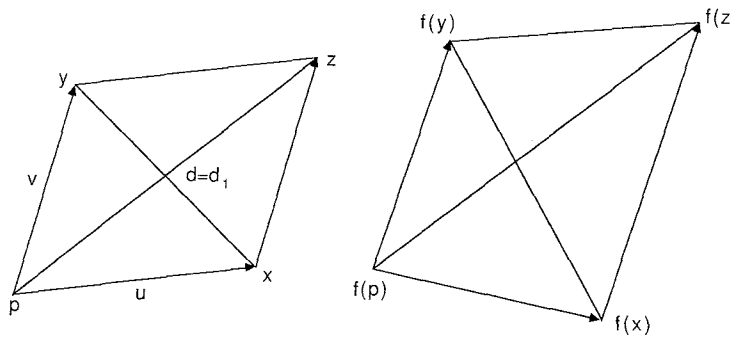
DEMOSTRACIÓ: Fixem $p \in A_1$, i definim

$$\begin{array}{rccc} \tilde{f}: & E_1 & \longrightarrow & E_2 \\ & u = \overrightarrow{px} & \longmapsto & \tilde{f}(u) = \overrightarrow{f(p)f(x)}. \end{array}$$

(Veure el començament de la demostració de (3.5)). Provem que $\tilde{f}$ és lineal. Siguin $u = \overrightarrow{px}$, $v = \overrightarrow{xz} = \overrightarrow{py}$ i $u + v = \overrightarrow{px} + \overrightarrow{xz} = \overrightarrow{pz}$. Designem per d el punt mig del parell p, z i per d_1 el del parell y, x. Aquests punts existeixen sempre que a K $2 \neq 0$ (IX.6). La condició $\overrightarrow{px} + \overrightarrow{py} = \overrightarrow{pz}$ equival a $d = d_1$; en efecte,

$$d = \tfrac{1}{2}p + \tfrac{1}{2}z \quad \Leftrightarrow \quad \overrightarrow{pd} = \tfrac{1}{2}\overrightarrow{pz}$$
$$d_1 = \tfrac{1}{2}y + \tfrac{1}{2}x \quad \Leftrightarrow \quad \overrightarrow{pd_1} = \tfrac{1}{2}\overrightarrow{py} + \tfrac{1}{2}\overrightarrow{px}.$$

Per tant, si $d = d_1$, les segones igualtats donen $\overrightarrow{pz} = \overrightarrow{py} + \overrightarrow{px}$. Recíprocament, si



$\overrightarrow{pz} = \overrightarrow{py} + \overrightarrow{px}$, obtenim $\overrightarrow{pd} = \overrightarrow{pd_1}$; és a dir, $d = d_1$.

Observem també que les expressions anteriors ens diuen que d és el punt mig de p, z si i només si $(pzd) = \frac{1}{2}$. Anàlogament, que d sigui el punt mig de y, x equival a $(yxd) = \frac{1}{2}$. Ara bé, per hipòtesi, f conserva les raons simples; per tant, $(f(p)f(z)f(d)) = \frac{1}{2}$ i $(f(y)f(x)f(d)) = \frac{1}{2}$, que equival al fet que $f(d)$ sigui el punt mig dels parells $f(p), f(z)$ i $f(y), f(x)$. Però més amunt ja hem vist que si aquests punts mitjos coincideixen $\overrightarrow{f(p)f(z)} = \overrightarrow{f(p)f(y)} + \overrightarrow{f(p)f(x)}$. Per la definició de $\tilde{f}$ això no és res més que

$$\tilde{f}(u + v) = \tilde{f}(v) + \tilde{f}(u).$$

Sigui ara $ku = k\overrightarrow{px} = \overrightarrow{py}$. Llavors,

$$(pxy) = k \quad \Rightarrow \quad (f(p)f(x)f(y)) = k \quad \Leftrightarrow$$
$$\Leftrightarrow \quad \overrightarrow{f(p)f(y)} = k\,\overrightarrow{f(p)f(x)} \quad \Leftrightarrow \quad \tilde{f}(ku) = k\tilde{f}(u).$$

Això acaba la demostració que $\tilde{f}$ és lineal. Per veure que $\tilde{f}$ és precisament l'aplicació lineal associada a f procedim com a la demostració de (3.5). $\forall a, b \in A$

$$\tilde{f}(\overrightarrow{ab}) = \tilde{f}(\overrightarrow{pb} - \overrightarrow{pa}) = \tilde{f}(\overrightarrow{pb}) - \tilde{f}(\overrightarrow{pa}) = \overrightarrow{f(p)f(b)} - \overrightarrow{f(p)f(a)} = \overrightarrow{f(a)f(b)}.$$

Així doncs, f és una afinitat.□

X.4 Equacions d'una afinitat en una referència cartesiana

Sigui $f : A_1 \longrightarrow A_2$ una afinitat. Considerem sistemes de referència cartesians $\{p;\ e_1,\ldots,e_n\}$, $\{q;\ u_1,\ldots,u_m\}$ dels espais A_1, A_2 respectivament. Sabem que $\forall x \in A_1$

$$f(x) = f(p) + \tilde{f}(\overrightarrow{px}).$$

Si les coordenades de x en el sistema $\{p;\ e_1,\ldots,e_n\}$ són $(x^1,\ldots,x^n)$ i $M = (a_i^j)$ és la matriu de $\tilde{f}$ en les bases $\{e_1,\ldots,e_n\}$ i $\{u_1,\ldots,u_m\}$ dels espais vectorial associats E_1, E_2, les coordenades del vector $\tilde{f}(\overrightarrow{px})$ són els termes de la matriu-columna

$$Mx \qquad \text{on} \qquad x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

Siguin ara $(b^1,\ldots,b^m)$ les coordenades del punt $f(p)$ en el sistema de referència $\{q;\ u_1,\ldots,u_m\}$, i indiquem per b la matriu-columna formada per aquestes coordenades. Aleshores, clarament, les coordenades $(\bar{x}^1,\ldots,\bar{x}^m)$ de $f(x)$ són els elements de la matriu-columna

$$\bar{x} = b + Mx.$$

Aquesta expressió s'escriu sovint de la següent forma: siguin

$$N = \begin{pmatrix} a_1^1 & \ldots & a_n^1 & b^1 \\ \vdots & & \vdots & \vdots \\ a_1^m & \ldots & a_n^m & b^m \\ 0 & \ldots & 0 & 1 \end{pmatrix} = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} \bar{x}^1 \\ \vdots \\ \bar{x}^m \\ 1 \end{pmatrix} = \begin{pmatrix} \bar{x} \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} x^1 \\ \vdots \\ x^n \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ 1 \end{pmatrix}.$$

Aleshores,

$$\begin{pmatrix} \bar{x} \\ 1 \end{pmatrix} = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}.$$

La matriu $\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$ es diu la *matriu de l'afinitat* f en els sistemes de referència $\{p;\ e_1,\ldots,e_n\}$, $\{q;\ u_1,\ldots,u_m\}$. L'expressió desenvolupada

$$\begin{cases} \bar{x}^1 = a_1^1 x^1 + \cdots + a_n^1 x^n + b^1 \\ \cdots\ \cdots\ \cdots \\ \bar{x}^m = a_1^m x^1 + \cdots + a_n^m x^n + b^m \end{cases}$$

són les *equacions de l'afinitat* f en els sistemes anteriors.

Observació:

Com acabem de veure, fixats sistemes de referència, tota afinitat té unes equacions lineals que permeten calcular les coordenades de la imatge d'un punt $f(x)$ a partir de les coordenades del punt x. Recíprocament, tota aplicació $g : A_1 \longrightarrow A_2$ donada per unes equacions d'aquest tipus

$$\begin{cases} \bar{x}^1 = c_1^1 x^1 + \cdots + c_n^1 x^n + d^1 \\ \cdots\cdots\cdots \\ \bar{x}^m = c_1^m x^1 + \cdots + c_n^m x^n + d^m \end{cases}$$

és una afinitat. En efecte, considerem l'aplicació lineal $\varphi : E_1 \longrightarrow E_2$ que té per matriu $C = (c_i^j)$ en les bases corresponents als sistemes de referència fixats. Si $(x^1, \ldots, x^n)$ són les coordenades de $x \in A_1$, les coordenades de $\varphi(\overrightarrow{px})$ són

$$Cx = \begin{pmatrix} c_1^1 x^1 + \cdots + c_n^1 x^n \\ \vdots \\ c_1^m x^1 + \cdots + c_n^m x^n \end{pmatrix}.$$

La imatge del punt $p = (0, \ldots, 0)$ per g és el punt de coordenades

$$d = \begin{pmatrix} d^1 \\ \vdots \\ d^m \end{pmatrix}.$$

Per tant, la imatge de qualsevol punt x té per coordenades $Cx + d$. És a dir,

$$g(x) = g(p) + \varphi(\overrightarrow{px}).$$

Això ens diu que g és una afinitat amb aplicació lineal associada φ.

Aquests arguments proven també que la matriu i les equacions d'una afinitat en uns sistemes de referència fixats estan unívocament determinades.

Proposició 4.1 *Siguin* $N = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$, $R = \begin{pmatrix} C & d \\ 0 & 1 \end{pmatrix}$ *les matrius de les afinitats* $f : A_1 \longrightarrow A_2$, $g : A_2 \longrightarrow A_3$ *en uns certs sistemes de referència de* A_1, A_2, A_3. *Aleshores* RN *és la matriu de l'afinitat* $g \circ f$.

DEMOSTRACIÓ: $\forall x \in A_1$ tenim, amb les notacions usuals,

$$\begin{pmatrix} g(f(x)) \\ 1 \end{pmatrix} = R \begin{pmatrix} f(x) \\ 1 \end{pmatrix} = RN \begin{pmatrix} x \\ 1 \end{pmatrix},$$

on

$$RN = \begin{pmatrix} CM & Cb + d \\ 0 & 1 \end{pmatrix}.$$

L'observació feta més amunt ens diu que $g \circ f$ és una afinitat amb matriu RN.□

Corol·lari 4.2 *Si $f : A \longrightarrow A$ és una afinitat bijectiva amb matriu N, f^{-1} és una afinitat bijectiva amb matriu N^{-1}*□

Considerem ara el cas particular de l'aplicació identitat

$$I : A \longrightarrow A.$$

Si fixem el mateix sistema de referència $\{p;\ e_1, \ldots, e_n\}$ per escriure les coordenades dels punts $x \in A$ i de les seves imatges, la matriu de I és, clarament, la matriu identitat. Si, al contrari, considerem les coordenades dels punts $x \in A$ en un sistema $\{q;\ v_1, \ldots, v_n\}$ i les de les seves imatges $f(x)$ en un altre sistema $\{p;\ e_1, \ldots, e_n\}$, obtenim una matriu

$$\begin{pmatrix} V & q \\ 0 & 1 \end{pmatrix}$$

on $V = (v_i^j)$ és la matriu de $\tilde{I}$:

$$\tilde{I}(v_i) = v_i = \sum_{j=1}^{n} v_i^j e_j$$

i $q = \begin{pmatrix} q^1 \\ \vdots \\ q^n \end{pmatrix}$ són les coordenades de $I(q) = q$ en el sistema $\{p;\ e_1, \ldots, e_n\}$. Obtenim així

$$\begin{pmatrix} \bar{x} \\ 1 \end{pmatrix} = \begin{pmatrix} V & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix},$$

que ens dóna les coordenades $\bar{x}$ de $I(x) = x$ en el sistema $\{p;\ e_1, \ldots, e_n\}$ a partir de les coordenades x del punt $x \in A$ en el sistema $\{q;\ v_1, \ldots, v_n\}$. Aquest resultat l'havíem obtingut ja a (IX.8).

Suposem ara que

$$N = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$$

és la matriu de $f : A_1 \longrightarrow A_2$ en els sistemes de referència $\{p;\ e_1, \ldots, e_n\}$ i $\{q;\ u_1, \ldots, u_m\}$ de A_1 i A_2 respectivament. Quina és, aleshores, la matriu de f en uns altres sistemes de referència $\{p_1;\ v_1, \ldots, v_n\}$, $\{q_1;\ w_1, \ldots, w_m\}$? Considerem f descomposta en la forma

$$f = I_{A_2} \circ f \circ I_{A_1} : A_1 \longrightarrow A_1 \longrightarrow A_2 \longrightarrow A_2,$$

on I_{A_i} és la identitat de A_i, $i = 1, 2$. A cada un d'aquests quatre espais considerem respectivament els sistemes

$$\{p_1;\ v_1, \ldots, v_n\} \qquad \{p;\ e_1, \ldots, e_n\} \qquad \{q;\ u_1, \ldots, u_m\} \qquad \{q_1;\ w_1, \ldots, w_m\}.$$

Si $\overrightarrow{pp_1} = c^1e_1 + \cdots + c^ne_n$ i $v_i = \sum_{j=1}^{n} r_i^j e_j$, $i = 1, \ldots, n$, la matriu de I_{A_1} és

$$\begin{pmatrix} R & c \\ 0 & 1 \end{pmatrix}.$$

Si $\overrightarrow{qq_1} = d^1u_1 + \cdots + d^mu_m$ i $w_i = \sum_{j=1}^{m} s_i^j u_j$, $i = 1, \ldots, m$, la matriu de I_{A_2} és

$$\begin{pmatrix} S & d \\ 0 & 1 \end{pmatrix}^{-1}.$$

La matriu de f en els sistemes $\{p_1;\ v_1, \ldots, v_n\}$, $\{q_1;\ w_1, \ldots, w_n\}$ és doncs

$$\begin{pmatrix} S & d \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R & c \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} S^{-1}MR & e \\ 0 & 1 \end{pmatrix},$$

on $e = S^{-1}(Mc + b - d)$.

X.5 El grup afí

Donat un espai afí A de dimensió n, denotarem per $GA(A)$ o simplement $GA(n)$ el grup de les aplicacions afins bijectives de A en A amb la composició.

Exercici:

Demostreu que si $A_1 \cong A_2$ com a espais afins llavors $GA(A_1)$ i $GA(A_2)$ són isomorfs com a grups.

Aquest fet justifica la notació $GA(n)$, ja que tots els espais afins de dimensió n donen lloc a grups isomorfs.

Considerem l'aplicació

$$\begin{array}{rccc} \Phi: & GA(n) & \longrightarrow & GL(n) \\ & f & \longmapsto & \tilde{f} \end{array}$$

on $GL(n)$ és el *grup lineal d'ordre n*, és a dir, el conjunt d'automorfismes de E amb la composició. Per (1.3) Φ és exhaustiva i per (2.1) el seu nucli està format per les translacions $\mathcal{T}$. A més a més, Φ és un morfisme de grups. El teorema d'isomorfisme de grups (III.5.2) ens diu llavors que

$$GA(n)/\mathcal{T} \cong GL(n).$$

A cada classe d'equivalència respecte a $\mathcal{T}$ hi ha totes les afinitats amb la mateixa aplicació lineal associada. Així, si

$$\begin{cases} \bar{x}^1 = a_1^1 x^1 + \cdots + a_n^1 x^n + b^1 \\ \qquad \cdots\ \cdots\ \cdots \\ \bar{x}^n = a_1^n x^1 + \cdots + a_n^n x^n + b^n \end{cases}$$

són les equacions d'una afinitat $f : A \longrightarrow A$ en un determinat sistema de referència, la classe de f està formada per totes les afinitats amb equacions que difereixen d'aquestes només en els "termes independents" $b^1, \ldots, b^n$. En particular, una afinitat de la classe és l'afinitat g donada per

$$\begin{cases} \bar{x}^1 = a_1^1 x^1 + \cdots + a_n^1 x^n \\ \qquad \cdots\ \cdots\ \cdots \\ \bar{x}^n = a_1^n x^1 + \cdots + a_n^n x^n \ , \end{cases}$$

que transforma el punt $(0, \ldots, 0)$ en el punt $(0, \ldots, 0)$. És a dir, si hem escollit un únic sistema de referència per escriure tots els punts, aquesta afinitat g deixa fix l'origen del sistema. L'afinitat original f s'obté component g amb una translació de vector $v = (b^1, \ldots, b^n)$

$$f = T_v \circ g.$$

Observem, però, que $f \neq g \circ T_v$.

Exemple:

Sigui A un espai afí real tridimensional i siguin

$$\begin{cases} \bar{x} = x + 2y + z + 1 \\ \bar{y} = y - z + 2 \\ \bar{z} = x + 3z - 1 \end{cases}$$

les equacions d'una afinitat $f : A \longrightarrow A$ en un cert sistema de referència $\{p;\ e_1, e_2, e_3\}$.

Estudiem primer $\tilde{f} : E \longrightarrow E$. La seva matriu és

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 3 \end{pmatrix}.$$

$\tilde{f}$ té tres valors propis 0, 2, 3 i els seus subespais de vectors propis són respectivament

$$E_0 = \operatorname{Nuc} \tilde{f} = \langle (-3, 1, 1) \rangle \quad E_2 = \langle (1, 1, -1) \rangle \quad E_3 = \langle (0, 1, -2) \rangle.$$

La imatge d'un vector $u \in E$ s'obté fàcilment descomponent-lo en suma de vectors de E_0, E_2 i E_3: $u = u_0 + u_2 + u_3$. Llavors $\tilde{f}(u) = 2u_2 + 3u_3$. En particular, observem que $\operatorname{Im} \tilde{f} = E_2 \oplus E_3$.

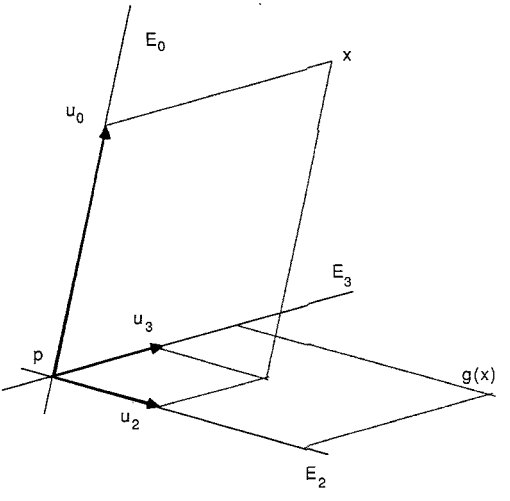
Suposem ara que $g : A \longrightarrow A$ és l'afinitat d'equacions

$$\begin{cases} \bar{x} = x + 2y + z \\ \bar{y} = y - z \\ \bar{z} = x + 3z. \end{cases}$$

És a dir, $\tilde{g} = \tilde{f}$ i $g(p) = p$. Aleshores $\forall x \in A$

$$g(x) = g(p) + \tilde{g}(\overrightarrow{px}) = p + \tilde{g}(\overrightarrow{px})$$

i si $\overrightarrow{px} = u_0 + u_2 + u_3$ amb $u_i \in E_i$, $i = 0, 2, 3$,



$$g(x) = p + 2u_2 + 3u_3. \qquad \text{(Veure la figura).}$$

L'afinitat inicial f s'obté component g amb la translació de vector $(1, 2, -1) = f(p)$.

Estudiem els punts fixos de f. Seran les solucions del sistema

$$\begin{cases} x = x + 2y + z + 1 \\ y = y - z + 2 \\ z = x + 3z - 1. \end{cases}$$

Existeix, per tant, un únic punt fix que és $q = (-3, -3/2, 2)$. Aleshores, $\forall x \in A$

$$f(x) = f(q) + \tilde{f}(\overrightarrow{qx}) = q + \tilde{f}(\overrightarrow{qx})$$

i si $\overrightarrow{qx} = u_0 + u_2 + u_3$ amb $u_i \in E_i$, $i = 0, 2, 3$,

$$f(x) = q + 2u_2 + 3u_3.$$

La situació és doncs la mateixa que per a g substituint p per q.

X.6 Varietats invariants

Donada una afinitat $f : A \longrightarrow A$ d'un espai afí (A,E) en ell mateix, direm que la varietat lineal $q + F$ és *invariant* o *doble* per f si

$$f(q + F) \subset q + F.$$

Per (3.1), $f(q+F) = f(q)+\tilde{f}(F)$. Resulta, doncs, que la varietat $q+F$ és invariant si i només si

1. $\tilde{f}(F) \subset F$.
2. $\overrightarrow{qf(q)} \in F$.

La primera condició ens diu que les direccions de les varietats dobles són subespais invariants. En particular, si la varietat és una recta i $F = \langle u \rangle$, és necessari que u sigui un vector propi perquè la recta pugui ésser doble.

Exemple:

Estudiem les rectes invariants per l'afinitat

$$\begin{cases} \bar{x} = -2y + 1 \\ \bar{y} = x + 3y - 1. \end{cases}$$

L'aplicació lineal associada té dos valors propis: 1 i 2. Els subespais de vectors propis són $E_1 = \langle (2, -1) \rangle$, $E_2 = \langle (1, -1) \rangle$. Qualsevol recta invariant ha de tenir com a direcció un d'aquests dos subespais. A més a més, si q és un punt de la recta, $\overrightarrow{qf(q)} \in E_i$, $i = 1$ o 2. Si les coordenades de q són (x_0, y_0) les de $\overrightarrow{qf(q)}$ són

$$(\bar{x}_0 - x_0, \bar{y}_0 - y_0) = (-x_0 - 2y_0 + 1, x_0 + 2y_0 - 1).$$

Aquest vector sempre és de E_2. Per tant, totes les rectes amb aquesta direcció són invariants. En canvi, $\overrightarrow{qf(q)}$ és de E_1 només si $x_0 + 2y_0 - 1 = 0$. És a dir, l'única recta invariant amb direcció E_1 és la recta

$$x + 2y - 1 = 0.$$

Un punt és una varietat lineal de dimensió 0, i és "invariant" si i només si es transforma en ell mateix; és a dir, si és un punt fix. Els punts fixos d'una afinitat f d'equació

$$\bar{x} = Mx + b$$

compleixen $x = Mx + b$; és a dir,

$$(M - I)x + b = 0.$$

Aquest sistema d'equacions representa la varietat de punts fixos.

Exemple:

La varietat de punts fixos de l'afinitat de l'exemple anterior està donada pel sistema

$$\begin{cases} -x-2y+1=0 \\ x+2y-1=0\ . \end{cases}$$

Es tracta, doncs, de la recta $x+2y-1=0$. Observem que, naturalment, aquesta és una de les rectes invariants de l'afinitat.

El sistema $(M-I)x+b=0$, que dóna els punts fixos, té solució única si i només si $\det(M-I)\neq 0$. Això demostra la següent proposició:

Proposició 6.1 *Una afinitat té un únic punt fix si i només si l'aplicació lineal associada no té el valor propi 1.*□

Acabarem aquest apartat donant un mètode per a calcular els hiperplans invariants per una afinitat bijectiva f d'equacions

$$\begin{cases} \bar{x}^1 = a_1^1x^1+\cdots+a_n^1x^n+b^1 \\ \qquad\cdots\ \cdots\ \cdots \\ \bar{x}^n = a_1^nx^1+\cdots+a_n^nx^n+b^n\ . \end{cases}$$

Considerem un hiperplà H d'equació $c_1x^1+\cdots+c_nx^n+c=0$. Un punt $(x^1,\ldots,x^n)$ s'aplica en H si i només si

$$0=c_1\bar{x}^1+\cdots+c_n\bar{x}^n+c=c_1(a_1^1x^1+\cdots+a_n^1x^n+b^1)+\cdots$$
$$\cdots+c_n(a_1^nx^1+\cdots+a_n^nx^n+b^n)+c=$$
$$=(c_1a_1^1+\cdots+c_na_1^n)x^1+\cdots+(c_1a_n^1+\cdots+c_na_n^n)x^n+c_1b^1+\cdots+c_nb^n+c.$$

Aquesta és, doncs, l'equació de l'hiperplà $f^{-1}(H)$. H és invariant quan $H=f^{-1}(H)$; les equacions de H i $f^{-1}(H)$ han de tenir, per tant, coeficients proporcionals:

$$\frac{c_1}{c_1a_1^1+\cdots+c_na_1^n}=\cdots=\frac{c_n}{c_1a_n^1+\cdots+c_na_n^n}=\frac{c}{c_1b^1+\cdots+c_nb^n+c}\ .$$

Són hiperplans invariants per f tots els que tenen equacions amb coeficients que compleixen aquestes igualtats.

Exemple:

Tornem a calcular les rectes invariants de l'afinitat del primer exemple d'aquest apartat:

$$\begin{cases} \bar{x}=-2y+1 \\ \bar{y}=x+3y-1. \end{cases}$$

Una recta $ax+by+c=0$ serà invariant si coincideix amb la d'equació

$$0=a\bar{x}+b\bar{y}+c=a(-2y+1)+b(x+3y-1)+c=$$
$$=bx+(-2a+3b)y+a-b+c.$$

És a dir, si

$$\frac{a}{b} = \frac{b}{-2a+3b} = \frac{c}{a-b+c}.$$

Clarament $a = 0$ si i només si $b = 0$. Però a i b no poden ésser simultàniament zero; per tant, cap d'elles no ho és. Posem $k = \frac{a}{b}$. La primera igualtat dóna:

$$k = \frac{1}{-2k+3} \quad \Leftrightarrow \quad 2k^2 - 3k + 1 = 0 \quad \Leftrightarrow \quad k = 1, \frac{1}{2}.$$

Si $k = 1$, $a = b$ i les dues igualtats es compleixen per a tot c. Així doncs totes les rectes $ax + ay + c = 0$ són invariants. Si $k = \frac{1}{2}$, $b = 2a$ i, per tant, $c = -a$. La recta $ax + 2ay - a = 0$, $a \neq 0$, és invariant.

X.7 Classificació de les afinitats d'un espai afí A en ell mateix

Hi ha moltes maneres de classificar les afinitats i cada una d'elles correspon a una relació d'equivalència (I.4). En aquest capítol anem a classificar-les de la següent forma: dues afinitats $f, g : A \longrightarrow A$ són de la *mateixa classe* si i només si existeix un isomorfisme d'espais afins $\varphi : A \longrightarrow A$ tal que el diagrama

$$\begin{array}{ccc} A & \xrightarrow{f} & A \\ \varphi\downarrow & & \downarrow\varphi \\ A & \xrightarrow{g} & A \end{array}$$

és commutatiu; és a dir, $\varphi \circ f = g \circ \varphi$. És fàcil veure que aquesta relació és d'equivalència.

Les propietats comunes a totes les afinitats d'una mateixa classe es diuen *propietats afins* de l'afinitat. Són propietats afins, per exemple, la dimensió de la varietat imatge, la dimensió de la varietat de punts fixos, ésser una simetria ($f^2 = I$), ésser una projecció ($f^2 = f$), etc. També són invariants dins de cada classe els valors de $\det \vec{f}$ i traça $\vec{f}$, ja que si $\varphi \circ f = g \circ \varphi$ llavors $\vec{g} = \vec{\varphi} \circ \vec{f} \circ \vec{\varphi}^{-1}$.

Proposició 7.1 *Dues afinitats $f, g : A \longrightarrow A$ són de la mateixa classe si i només si existeixen dos sistemes de referència tals que les equacions de f en un d'ells, $\{p;\ e_1, \ldots, e_n\}$, coincideixen amb les equacions de g en l'altre, $\{q;\ v_1, \ldots, v_n\}$.*

DEMOSTRACIÓ: Suposem que f i g són de la mateixa classe i sigui $\varphi : A \longrightarrow A$ un isomorfisme tal que $\varphi \circ f = g \circ \varphi$. Considerem un sistema de referència qualsevol de A, $\{p;\ e_1, \ldots, e_n\}$, i sigui N la matriu de f en aquest sistema. Per ésser $\vec{\varphi}$ un isomorfisme d'espais vectorials, $\vec{\varphi}(e_1), \ldots, \vec{\varphi}(e_n)$ és una base. La matriu de φ considerant en el primer espai el sistema $\{p;\ e_1, \ldots, e_n\}$ i en el segon espai el sistema $\{\varphi(p);\ \vec{\varphi}(e_1), \ldots, \vec{\varphi}(e_n)\}$ és, clarament, la matriu identitat I. Aleshores, si R és la

matriu de g en el sistema $\{\varphi(p);\ \tilde{\varphi}(e_1),\ldots,\tilde{\varphi}(e_n)\}$, de la igualtat $\varphi \circ f = g \circ \varphi$ resulta

$$N = IN = RI = R.$$

Això demostra una implicació.

Suposem ara que N és la matriu de f en el sistema $\{p;\ e_1,\ldots,e_n\}$ i també la matriu de g en un altre sistema $\{q;\ v_1,\ldots,v_n\}$. Considerem l'afinitat $\varphi : A \longrightarrow A$ tal que $\varphi(p) = q$, $\tilde{\varphi}(e_i) = v_i$, $i = 1,\ldots,n$. Clarament, φ és un isomorfisme i la seva matriu, agafant el sistema $\{p;\ e_1,\ldots,e_n\}$ en el primer espai i el sistema $\{q;\ v_1,\ldots,v_n\}$ en el segon, és la matriu identitat I. La igualtat $\varphi \circ f = g \circ \varphi$ es dedueix aleshores de $IN = NI$. □

Exemple:

Dues translacions T_u, T_v de vectors no zero són sempre de la mateixa classe. Per demostrar-ho necessitem un isomorfisme tal que $\varphi \circ T_u(a) = T_v \circ \varphi(a), \forall a$. És a dir, tal que $\varphi(a) + \tilde{\varphi}(u) = \varphi(a) + v$. N'hi ha prou doncs amb escollir un isomorfisme d'espais vectorials $\tilde{\varphi}$ que transformi u en v, i qualsevol afinitat φ amb aquesta aplicació lineal associada serà un isomorfisme com el buscat.

Les equacions de T_u, $u \neq \vec{0}$, en una referència del tipus $\{p;\ u, e_2,\ldots,e_n\}$ són

$$\begin{cases} \bar{x}^1 = x^1 + 1 \\ \bar{x}^i = x^i \qquad i = 2,\ldots,n. \end{cases}$$

Qualsevol translació, en un sistema de referència convenient, té aquestes equacions.

En els apartats següents classificarem les afinitats de la recta i del pla afí. Per fer-ho ens basarem en la proposició (7.1): buscarem els diferents tipus d'equacions que poden tenir les afinitats en cada cas. Comprovarem que no es pot passar d'un d'aquests tipus a un altre per un canvi del sistema de referència; és a dir, que corresponen a afinitats de classe diferent. Llavors cada un dels tipus d'equacions correspondrà a una classe d'afinitats.

X.8 Afinitats de la recta afí

Sigui A un espai afí de dimensió 1, i $f : A \longrightarrow A$ una afinitat. $\tilde{f}$ és un endomorfisme d'un espai vectorial de dimensió 1 i, per tant (V.5.1), és de la forma $\tilde{f} = aI$. En qualsevol base la matriu de $\tilde{f}$ és (a) i $\det \tilde{f} = a$. D'aquí que afinitats amb paràmetres a diferents siguin de classes diferents.

Examinem els casos possibles. Si $a = 0$, $\tilde{f} = 0$ i f transforma tot A en un punt:

$$f(a) = f(p) + \tilde{f}(\overrightarrow{pa}) = f(p) \qquad \forall a.$$

Agafant com a origen del sistema de referència aquest punt $q = f(a)$ i una base qualsevol de l'espai vectorial associat, obtenim l'equació de f

$$\bar{x} = 0.$$

Si $a = 1$, $\tilde{f} = I$ i, per tant, f és una translació. Si el vector translació és $u \neq \vec{0}$, en un sistema $\{p;\ u\}$ (p qualsevol), l'equació de f és

$$\bar{x} = x + 1.$$

Si el vector translació és $\vec{0}$, f és la identitat i, en qualsevol sistema de referència, la seva equació és

$$\bar{x} = x.$$

Si $a \neq 1$, f és una homotècia de raó a i té només un punt fix (§2). Agafant aquest punt com a origen i una base qualsevol de l'espai vectorial associat, l'equació de f és

$$\bar{x} = ax.$$

Observem que hem obtingut una classe per a cada valor de $a = \det \tilde{f}$, llevat del cas $a = 1$ que correspon a dues classes diferents.

X.9 Afinitats del pla afí

Sigui A un espai afí de dimensió 2. Farem l'estudi de les classes d'afinitats de A en tres etapes:

- Afinitats amb una recta de punts fixos.
- Afinitats sense cap punt fix.
- Afinitats amb un únic punt fix.

Suposem, primer, que $f : A \longrightarrow A$ té una recta $q + \langle u \rangle$ de punts fixos. En un sistema de referència $\{q;\ u, v\}$ (v qualsevol que formi base amb u), les equacions de f són

$$\begin{cases} \bar{x} = x + by \\ \bar{y} = ny. \end{cases}$$

Si $n \neq 1$, n és un altre valor propi de $\tilde{f}$ i podem escollir com a segon vector de la base un vector v de valor propi n. Les equacions de f són llavors

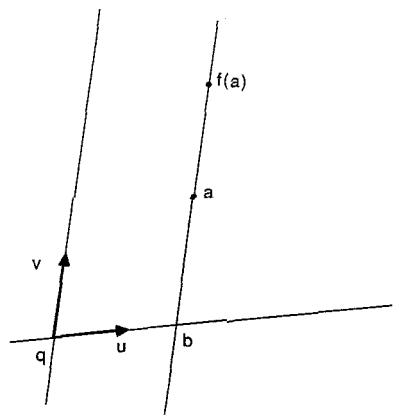
$$\begin{cases} \bar{x} = x \\ \bar{y} = ny. \end{cases}$$

L'afinitat es diu aleshores una *homologia general de raó* n. Observem que $n = \det \tilde{f}$, i, per tant, homologies generals de raons diferents pertanyen a classes diferents. És

fàcil veure que una homologia general f deixa invariants totes les rectes de direcció $\langle v\rangle$ i la recta de punts fixos $q+\langle u\rangle$. Per altra banda, la imatge d'un punt $a=(x,y)$ és $f(a)=(\bar{x},\bar{y})=(x,ny)$. La recta determinada per a i $f(a)$ s'interseca amb la recta de punts fixos en un punt $b=(x,0)$, que compleix $\overrightarrow{bf(a)}=n\,\overrightarrow{ba}$. És a dir,

$$(b,a,f(a))=n.$$

Considerem ara el cas en què $n=1$. L'únic valor propi de l'aplicació lineal associada



$\tilde{f}$ és 1. Si aquest valor propi té multiplicitat 2, $\tilde{f}=I$ i f és una translació amb una recta de punts fixos; és a dir, $f=I$. Les seves equacions en qualsevol sistema són:

$$\begin{cases}\bar{x}=x\\ \bar{y}=y.\end{cases}$$

Si la multiplicitat del valor propi 1 és 1, les equacions de f en el sistema $\{q;\ u,v\}$ són

$$\begin{cases}\bar{x}=x+by\\ \bar{y}=y\end{cases}$$

amb $b\neq 0$. Per tant, en la referència $\{q;\ bu,v\}$, les equacions de f són

$$\begin{cases}\bar{x}=x+y\\ \bar{y}=y.\end{cases}$$

L'afinitat es diu llavors una *homologia especial*. Totes les homologies especials pertanyen a la mateixa classe. Les seves rectes invariants són totes les de direcció $\langle u\rangle$; una d'elles, $q+\langle u\rangle$, és recta de punts fixos.

Suposem ara que $f:A\longrightarrow A$ no té cap punt fix. $\tilde{f}$ té aleshores el valor propi 1 (6.1). Sigui $u\neq\vec{0}$ un vector propi de valor propi 1. Les equacions de f en un sistema $\{p;\ u,v\}$ (p i v qualssevol) són del tipus

$$\begin{cases}\bar{x}=x+by+c\\ \bar{y}=ny+d.\end{cases}$$

Si $n \neq 1$, n és un altre valor propi de $\tilde{f}$ i podem escollir, com a segon vector de la base, un vector de valor propi n. Les equacions seran llavors

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = ny + d \end{cases}$$

amb $c \neq 0$ perquè no hi hagi punts fixos. L'única recta invariant té equació

$$y = \frac{d}{1-n}.$$

Si escollim l'origen p del sistema de referència sobre aquesta recta, sigui $p = (x_0, \frac{d}{1-n})$, resulta que $\overrightarrow{pf(p)} = (c,0) = cu$. D'aquí resulta que en el sistema de referència $\{p;\ cu, v\}$ les equacions de f són

$$\begin{cases} \bar{x} = x + 1 \\ \bar{y} = ny, \end{cases}$$

$n = \det \tilde{f}$ i, per tant, hi ha tantes classes diferents com paràmetres $n \neq 1$. Aquestes afinitats són *homologies generals seguides d'una translació* de direcció la de la recta de punts fixos de l'homologia. Aquesta composició és a més a més commutativa.

Si $n = 1$, l'únic valor propi és 1. Si aquest és de multiplicitat 2, $\tilde{f} = I$ i l'afinitat és una translació de vector $w \neq \vec{0}$. En una referència del tipus $\{p;\ w, v\}$, les equacions de f són

$$\begin{cases} \bar{x} = x + 1 \\ \bar{y} = y. \end{cases}$$

Si la multiplicitat del valor propi 1 és 1, a les equacions de f en el sistema $\{p;\ u, v\}$ (u vector propi de valor propi 1)

$$\begin{cases} \bar{x} = x + by + c \\ \bar{y} = y + d \end{cases}$$

b ha d'ésser diferent de zero. A més a més, com que no existeixen punts fixos, $d \neq 0$. En aquestes circumstàncies l'afinitat no té tampoc cap recta invariant. Podem escollir, però, el vector $\overrightarrow{pf(p)} = cu + dv = w$ com a segon vector de la base; aleshores

$$\tilde{f}(w) = c\tilde{f}(u) + d\tilde{f}(v) = cu + d(bu + v) = dbu + w.$$

D'aquí resulta que en el sistema de referència $\{p;\ dbu, w\}$ les equacions de f són

$$\begin{cases} \bar{x} = x + y \\ \bar{y} = y + 1. \end{cases}$$

Es tracta, doncs, d'una *homologia especial seguida d'una translació* de direcció diferent de la del feix de rectes invariants.

Suposem, finalment, que $f : A \longrightarrow A$ té un únic punt fix q. Considerem dos casos:

1. Existeix un vector u tal que u, $\tilde{f}(u)$ són linealment independents. En el sistema de referència $\{q;\ u, \tilde{f}(u)\}$ les equacions de f són

$$\begin{cases} \bar{x} = by \\ \bar{y} = x + ny. \end{cases}$$

Observem que $n = \text{traça}\,\tilde{f}$, $b = -\det \tilde{f}$. Per tant, hi ha tantes classes d'afinitats d'aquest tipus com paràmetres b i n. Aquests paràmetres no poden, però, prendre tots els valors de K : la condició que f tingui un únic punt fix imposa $b + n \neq 1$.

El polinomi característic de $\tilde{f}$ és $x^2 - nx - b$. Anem a donar expressions més senzilles de les equacions de f en els casos en què $\tilde{f}$ té valors propis.

Si $\tilde{f}$ té dos valors propis diferents ($(n^2+4b) > 0$ quan $K = \mathbf{R}$), és diagonalitzable en una certa base $\{w, v\}$. Les equacions de f en un sistema $\{q;\ w, v\}$ són del tipus

$$\begin{cases} \bar{x} = ax \\ \bar{y} = cy \end{cases} \qquad a \neq c.$$

Si $n^2 + 4b = 0$, $\tilde{f}$ té un únic valor propi a. Sigui w un vector propi i $\{w, v\}$ una base. En el sistema $\{q;\ w, v\}$ les equacions de f són

$$\begin{cases} \bar{x} = ax + b'y \\ \bar{y} = n'y \end{cases}$$

amb $n' = a$ (en cas contrari n' seria un altre valor propi). A més a més, $b' \neq 0$, ja que en cas contrari $\tilde{f} = aI$ i tots els parells u, $\tilde{f}(u)$ serien linealment dependents. Podem, doncs, agafar el sistema de referència $\{q;\ b'w, v\}$ i obtenim com a equacions de f

$$\begin{cases} \bar{x} = ax + y \\ \bar{y} = ay \end{cases} \qquad a \neq 1.$$

2. Si per a tot vector u $\{u, \tilde{f}(u)\}$ són linealment dependents, $\tilde{f}$ ha d'ésser una homotècia vectorial. En efecte, sigui $\{u, v\}$ una base i $\tilde{f}(u) = au$, $\tilde{f}(v) = bv$. Cal doncs que $\tilde{f}(u + v) = au + bv = c(u + v)$, d'on $a = c = b$. Tots els vectors són, per tant, vectors propis del mateix valor propi i $\tilde{f} = aI$ ($a \neq 1$ perquè f té un únic punt fix). L'afinitat f és una *homotècia de raó* a (§2), i les seves equacions en el sistema $\{q;\ u, v\}$ són

$$\begin{cases} \bar{x} = ax \\ \bar{y} = ay. \end{cases}$$

En aquest cas, $a^2 = \det \tilde{f}$ i, per tant, afinitats amb paràmetre a diferent són de classes diferents.

X.10 Nota històrica

L'estudi de les transformacions adequades entre certes estructures adquireix tota la seva importància arran de la conferència que Felix Klein (1849-1925) va donar el 1872, amb motiu de la seva admissió a la Universitat d'Erlangen, amb el títol "Vergleichende Betrachtungen über neuere geometrische Forschungen" (Una revisió comparativa de recerques recents a geometria). Els punts de vista expressats en aquesta conferència es coneixen avui com el "Programa d'Erlangen".

La idea bàsica de Klein és que tota geometria pot ésser caracteritzada per un grup de transformacions i que la geometria tracta essencialment dels invariants per aquest grup de transformacions. La geometria afí queda caracteritzada pel grup de les afinitats (pel grup afí) i no és més que l'estudi dels invariants per aquest grup.

X.11 Exercicis

1. Sigui A un pla afí. Demostreu que donades dues rectes que es tallen i un punt que no pertany a cap de les dues rectes, i donada una altra configuració anàloga, existeixen dos afinitats de A que transformen una configuració en l'altra. Trobeu aquestes afinitats en el cas

$$r : x - y = 2, \quad s : x - 2y = -1, \quad p = (0,0)$$
$$r' : x = 1, \quad s' : x - y = 1, \quad p' = (2,2).$$

2. Escriviu l'equació de totes les homologies generals de $\mathbf{R}^2$ que tenen l'eix d'homologia paral·lel a l'eix d'abscisses.

3. Equacions de les afinitats del pla que transformen les rectes r_1, r_2, r_3 en r_2, r_3, r_1 respectivament, on

$$r_1 : x + y = 1, \quad r_2 : x + 2y = 0, \quad r_3 : 4x - y = 2.$$

Classifiqueu aquestes afinitats.

4. Estudieu les afinitats de $\mathbf{R}^2$ que deixen fixa la hipèrbola $xy = 1$.

5. Sigui Z la unió de dues rectes del pla afí A que es tallen. Descriviu el grup de les afinitats bijectives que deixen Z fix. Expliciteu aquest grup en el cas de les rectes de l'exercici 1.

6. Estudieu el grup d'afinitats del pla que deixen fix un triangle donat.

7. Dibuixeu la imatge dels punts $(0,0), (1,0), (1,1), (0,1)$ per a cada una de les formes simplificades de les afinitats del pla afí que s'han obtingut, prenent $K = \mathbf{R}$. Discutiu diferents valors dels paràmetres quan n'hi hagi.

8. Determineu el lloc geomètric de les imatges d'un punt donat x per totes les afinitats que tenen una recta donada r de punts fixos i una recta donada s, que es creua amb r, fixa.

9. Demostreu que hi ha una única afinitat del pla que transforma cadascun dels vèrtexs d'un triangle donat en el punt mig del costat oposat. Estudieu aquesta afinitat.

10. Sigui f una afinitat d'un espai afí real. Demostreu:
a) Si f^2 té algun punt fix, f també.
b) Si existeix un $n \in \mathbf{N}$ tal que f^n té algun punt fix, aleshores f també.

11. Sigui f una afinitat i $\tilde{f}$ l'endomorfisme associat. Demostreu:
a) $e \in \operatorname{Nuc} \tilde{f}$ si i només si totes les rectes de direcció $\langle e \rangle$ es transformen per f en un punt.
b) e és un vector propi de $\tilde{f}$ de valor propi diferent de zero si i només si totes les rectes de direcció $\langle e \rangle$ es transformen per f en una recta paral·lela.

12. Considerem tres rectes concurrents r, s, t del pla afí ordinari i dues rectes paral·leles l, l' que tallen r, s, t en els punts a, b, c i a', b', c' respectivament. Demostreu que $(abc) = (a'b'c')$.

13. A la família d'afinitats del pla d'equacions

$$\begin{cases} \bar{x} &= ax + ay + b \\ \bar{y} &= ax + 6y + b^2 \end{cases}$$

hi ha quatre homologies, els eixos de les quals són els costats d'un paral·lelogram. Determineu els vèrtexs d'aquest paral·lelogram.

14. Estudieu segons els valors del paràmetre a les afinitats donades per les equacions

$$\begin{cases} \bar{x} &= ax + y + z + 1 \\ \bar{y} &= x + ay + z + 1 \\ \bar{z} &= x + y + az + 1. \end{cases}$$

15. Estudieu l'afinitat d'equacions

$$\begin{cases} \bar{x} &= x - \frac{1}{8}y - \frac{1}{8} \\ \bar{y} &= 2x - \frac{1}{8} \end{cases}$$

i expresseu-la com a producte d'una homotècia i una homologia.

16. Estudieu totes les afinitats de la forma

$$\begin{cases} \bar{x} &= ax + y + a \\ \bar{y} &= x + ay + a. \end{cases}$$

Determineu el lloc geomètric de les imatges d'un punt donat per totes aquestes afinitats.

17. Estudieu les afinitats d'equacions

$$\begin{cases} \bar{x} &= (1+a)x - ay + 1 \\ \bar{y} &= a^2x + (1 + 2a - 4a^2 + a^3)y \end{cases}$$

que no tenen punts fixos.

18. Demostreu que un subconjunt de l'espai afí K^n és una varietat lineal de dimensió r si i només si és el conjunt de zeros d'una afinitat exhaustiva $f : K^n \longrightarrow K^{n-r}$.

X.12 Exercicis de programar

19. Feu un programa que canviï de sistemes de referència les equacions cartesianes d'una afinitat $f : \mathbf{R}^n \longrightarrow \mathbf{R}^m$. (Indicació: seguiu el mètode explicat al final del §4.)

20. Apliqueu l'exercici 19 per canviar de sistema de referència les coordenades dels punts de $\mathbf{R}^n$. (Indicació: considereu l'aplicació identitat $I : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ i les seves equacions cartesianes agafant sistemes de referència diferents a dreta i a esquerra).

21. Apliqueu l'exercici 19 per canviar de sistema de referència les equacions cartesianes d'una varietat lineal de $\mathbf{R}^n$. (Indicació: feu servir els exercicis IX.14 i X.18).

22. Feu un programa que permeti trobar les equacions de la varietat de punts fixos d'una afinitat $f : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ donada per les seves equacions cartesianes en una certa referència. (Indicació: si $f(x) = Ax + b$, reduïu pel mètode de Gauss el sistema d'equacions $(A - I)x + b = \vec{0}$).

23. Sigui $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ una afinitat donada per $f(x) = Ax + b$. Feu un programa que calculi:

a) La varietat de punts fixos de f (exercici 22).

b) Els valors propis i els vectors propis reals de A (exercici VIII.23).

Seguint la classificació feta al §9, podeu obtenir les equacions simplificades de f i la referència en què s'obtenen.

Capítol XI

Espais vectorials euclidians i unitaris

XI.1 Formes bilineals i sesquilineals

Sigui E un espai vectorial sobre $\mathbf{R}$. Una aplicació del producte cartesià $E \times E$ a $\mathbf{R}$

$$\phi : E \times E \longrightarrow \mathbf{R}$$

es diu una *forma bilineal* si compleix

i) $\phi(u_1 + u_2, v) = \phi(u_1, v) + \phi(u_2, v) \quad \forall u_1, u_2, v \in E$
$\phi(ku, v) = k\phi(u, v) \quad \forall u, v \in E \quad \forall k \in \mathbf{R}$;

ii) $\phi(u, v_1 + v_2) = \phi(u, v_1) + \phi(u, v_2) \quad \forall u, v_1, v_2 \in E$
$\phi(u, kv) = k\phi(u, v) \quad \forall u, v \in E \quad \forall k \in \mathbf{R}$.

Proposició 1.1 *Sigui $e_1, \ldots, e_n$ una base de l'espai vectorial real E.*

1. *La matriu $B = (b_i^j)$, on $b_i^j = \phi(e_j, e_i)$, determina ϕ. Més concretament, si $u = \sum_{i=1}^n u^i e_i$, $v = \sum_{i=1}^n v^i e_i$*

$$\phi(u, v) = u^t B v,$$

on $v = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$ *i* $u^t = (u^1 \ldots u^n)$ *és la transposada de* $u = \begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix}$

2. *Donada una matriu $B = (b_i^j)$ qualsevol, existeix sempre una forma bilineal ϕ tal que $\phi(e_j, e_i) = b_i^j$.*

DEMOSTRACIÓ:

1. Aplicant les condicions i) i ii) de forma bilineal obtenim

$$\begin{aligned}\phi(u,v) &= \phi\left(\sum_{i=1}^n u^i e_i, \sum_{j=1}^n v^j e_j\right) = \sum_{i=1}^n u^i \phi\left(e_i, \sum_{j=1}^n v^j e_j\right) = \\ &= \sum_{i=1}^n u^i \left(\sum_{j=1}^n v^j \phi(e_i, e_j)\right) = \sum_{i,j=1}^n u^i b_j^i v^j = u^t B v.\end{aligned}$$

2. Donada B definim $\phi : E \times E \longrightarrow \mathbf{R}$ per

$$\phi(u,v) = u^t B v$$

(amb les notacions de l'enunciat). És fàcil provar que ϕ és bilineal. □

La matriu B de (1.1) es diu *matriu de ϕ en la base $e_1, \ldots, e_n$*. Sigui C la matriu de ϕ en una altra base $u_1, \ldots, u_n$ i sigui $A = (a_i^j)$ la matriu del canvi de base:

$$u_i = \sum_{j=1}^n a_i^j e_j \quad i = 1, \ldots, n.$$

Per (1.1), $c_j^i = \phi(u_i, u_j) = a_i^t B a_j$ on a_i representa la matriu formada per la i-èssima columna de A (és a dir les coordenades de u_i en la base $(e_1, \ldots, e_n)$. D'aquí que

$$C = A^t B A.$$

Sigui E un espai vectorial sobre els complexos $\mathbf{C}$. Una aplicació $\phi : E \times E \longrightarrow \mathbf{C}$ es diu una *forma sesquilineal* si compleix

i) $\phi(u_1 + u_2, v) = \phi(u_1, v) + \phi(u_2, v)\ \forall u_1, u_2, v \in E$
$\phi(ku, v) = k\phi(u, v)\ \forall u, v \in E \quad \forall k \in \mathbf{C}$;

ii) $\phi(u, v_1 + v_2) = \phi(u, v_1) + \phi(u, v_2)\ \forall u, v_1, v_2 \in E$
$\phi(u, kv) = \bar{k}\phi(u, v)\ \forall u, v \in E \quad \forall k \in \mathbf{C}$, on $\bar{k}$ indica el conjugat de $k \in \mathbf{C}$.

Proposició 1.2 *Sigui $e_1, \ldots, e_n$ una base de E (espai vectorial sobre $\mathbf{C}$).*

1. *La matriu $B = (b_i^j)$, on $b_i^j = \phi(e_j, e_i)$, determina ϕ. Més concretament, si $u = \sum_{i=1}^n u^i e_i$, $v = \sum_{i=1}^n v^j e_i$*

$$\phi(u,v) = u^t B \bar{v},$$

on $\bar{v} = \begin{pmatrix} \bar{v}^1 \\ \vdots \\ \bar{v}^n \end{pmatrix}$ *i* $u^t = (u^1 \ldots u^n)$ *és la transposada de* $u = \begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix}$.

2. *Donada una matriu complexa $B = (b_i^j)$ qualsevol, existeix sempre una forma sesquilineal ϕ tal que $\phi(e_j, e_i) = b_i^j$.*

La demostració és una adaptació fàcil de la de (1.1).□

La matriu B de (1.2) es diu *matriu de ϕ en la base* $e_1, \ldots, e_n$. Sigui C la matriu de ϕ en una altra base $u_1, \ldots, u_n$ i sigui $A = (a_i^j)$ la matriu del canvi de base:

$$u_i = \sum_{j=1}^{n} a_i^j e_j \qquad i = 1, \ldots, n.$$

Com en el cas real, es dedueix que

$$C = A^t B \bar{A}.$$

Una forma bilineal en un espai vectorial real E, $\phi : E \times E \longrightarrow \mathbf{R}$, es diu *simètrica* si $\phi(u, v) = \phi(v, u)\ \forall u, v \in E$.

Proposició 1.3 *Una forma bilineal és simètrica si i només si la seva matriu és simètrica (una matriu B es diu* simètrica *si $B^t = B$).*

DEMOSTRACIÓ: Si ϕ és simètrica, $b_j^i = \phi(e_i, e_j) = \phi(e_j, e_i) = b_i^j$. Si B és simètrica, $\phi(u, v) = \sum_{i,j=1}^{n} u^i b_j^i v^j = \sum_{i,j=1}^{n} v^j b_i^j u^i = \phi(v, u)$. □

Una forma sesquilineal sobre un espai vectorial complex E, $\phi : E \times E \longrightarrow \mathbf{C}$, es diu *hermítica* si $\phi(u, v) = \overline{\phi(v, u)}\ \forall u, v \in E$.

Proposició 1.4 *Una forma sesquilineal és hermítica si i només si la seva matriu B és hermítica (B es diu* hermítica *si $B^t = \bar{B}$).*

DEMOSTRACIÓ: Es procedeix com en el cas real. □

XI.2 Producte escalar

Sigui E un espai vectorial real. Un *producte escalar* a E és una forma bilineal simètrica

$$\phi : E \times E \longrightarrow \mathbf{R}$$

que compleix

$$\begin{aligned} &\phi(u, u) \geq 0 && \forall u \in E; \\ &\phi(u, u) = 0 \ \Leftrightarrow && u = \vec{0}. \end{aligned}$$

Una forma que compleix aquestes dues propietats es diu *definida positiva.*

Sigui E un espai vectorial complex. Un *producte escalar* a E és una forma sesquilineal hermítica $\phi : E \times E \longrightarrow \mathbf{C}$ definida positiva; és a dir, tal que

$$\begin{aligned} &\phi(u, u) \geq 0 \,(\text{real positiu}) && \forall u \in E; \\ &\phi(u, u) = 0 \ \Leftrightarrow && u = \vec{0}. \end{aligned}$$

Sigui E un espai vectorial real o complex amb un producte escalar ϕ. Un vector u es diu *unitari* si

$$\phi(u,u) = 1.$$

Dos vectors $u, v \in E$ es diuen *ortogonals* si

$$\phi(u,v) = 0.$$

Observacions:

1. Si $u \neq 0$, $\dfrac{u}{\sqrt{\phi(u,u)}}$ és unitari. (Indiquem per $\sqrt{\phi(u,u)}$ la determinació positiva de l'arrel).
2. Si S és un conjunt de vectors diferents de $\vec{0}$ i ortogonals dos a dos, S és linealment independent. En efecte, si $\sum \lambda^i v_i = \vec{0}$ amb $v_i \in S$, per a cada v_k tenim

$$0 = \phi(\sum \lambda^i v_i, v_k) = \sum \lambda^i \phi(v_i, v_k) = \lambda^k \phi(v_k, v_k).$$

   Però $\phi(v_k, v_k) \neq 0$, ja que $v_k \neq \vec{0}$. Per tant, $\lambda^k = 0\ \forall k$. El nostre objectiu immediat és demostrar que sempre existeix una base, $u_1, \ldots, u_n$, en la qual la matriu del producte escalar és la matriu identitat; és a dir,

$$\phi(u_i, u_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j. \end{cases}$$

   Els vectors $u_1, \ldots, u_n$ són, doncs, unitaris i ortogonals dos a dos. Direm llavors que $u_1, \ldots, u_n$ és una *base ortonormal.*
3. Si $w = w^1 u_1 + \ldots + w^n u_n$ i $v = v^1 u_1 + \ldots + v^n u_n$, on $u_1, \ldots, u_n$ és una base ortonormal,

$$\begin{aligned} \phi(w,v) &= w^1 v^1 + \ldots + w^n v^n \quad \text{en el cas real;} \\ \phi(w,v) &= w^1 \bar{v}^1 + \ldots + w^n \bar{v}^n \quad \text{en el cas complex.} \end{aligned}$$

4. Les coordenades d'un vector $v = v^1 u_1 + \ldots + v^n u_n$ en una base ortonormal $u_1, \ldots, u_n$ són

$$v^i = \phi(v, u_i).$$

Proposició 2.1 *Sigui E un espai vectorial de dimensió finita n sobre $\mathbf{R}$ o $\mathbf{C}$, amb un producte escalar ϕ. Existeix sempre una base ortonormal de E.*

DEMOSTRACIÓ: Sigui $e_1, \ldots, e_n$ una base qualsevol de E. Considerem els subespais

$$E_1 = \langle e_1 \rangle \subset E_2 = \langle e_1, e_2 \rangle \subset \ldots \subset E_n = \langle e_1, \ldots, e_n \rangle = E.$$

- E_1 té una base ortonormal que és $u_1 = \dfrac{e_1}{\sqrt{\phi(e_1, e_1)}}$.

- Suposem que $u_1, \ldots, u_r$ és una base ortonormal de E_r. Construïm una base ortonormal de $E_{r+1} = \langle u_1, \ldots, u_r, e_{r+1} \rangle$ de la següent manera: considerem un vector de la forma

$$u'_{r+1} = e_{r+1} - (k^1 u_1 + \ldots + k^r u_r)$$

ortogonal a cada u_i, $\forall i = 1, \ldots, r$: $0 = \phi(u'_{r+1}, u_i) = \phi(e_{r+1}, u_i) - k^i$. Hem d'agafar, doncs, $k^i = \phi(e_{r+1}, u_i)$, $i = 1, \ldots, r$. L'observació 2, ens diu que $u_1, \ldots, u_r, u'_{r+1}$ són linealment independents i formen, per tant, una base de E_{r+1}. Posem, aleshores,

$$u_{r+1} = \frac{u'_{r+1}}{\sqrt{\phi(u'_{r+1}, u'_{r+1})}}$$

i $u_1, \ldots, u_r, u_{r+1}$ serà una base ortonormal de E_{r+1}. Per inducció obtenim, així, que $E_n = E$ té una base ortonormal.□

Aquest procés de construcció d'una base ortonormal es coneix amb el nom de *mètode de Gram-Schmidt.* La proposició següent és el recíproc de (2.1):

Proposició 2.2 *Si una forma bilineal o sesquilineal ϕ (sobre* $\mathbf{R}$ *o* $\mathbf{C}$ *respectivament) té la matriu identitat en una base $u_1, \ldots, u_n$, aleshores ϕ és un producte escalar.*

DEMOSTRACIÓ: Si $w = w^1 u_1 + \ldots + w^n u_n$ i $v = v^1 u_1 + \ldots + v^n u_n$ resulta

$$\begin{aligned} \phi(w, v) &= w^1 v^1 + \ldots + w^n v^n \quad \text{en el cas real;} \\ \phi(w, v) &= w^1 \bar{v}^1 + \ldots + w^n \bar{v}^n \quad \text{en el cas complex.} \end{aligned}$$

D'aquí resulten fàcilment les propietats que ha de complir un producte escalar. □

Nota:

A partir d'ara, si no indiquem el contrari, per comoditat de notació, pensarem sempre que els escalars són complexos ($\mathbf{R} \subset \mathbf{C}$), i escriurem $\bar{k}$ allà on calgui, ben entès que $\bar{k} = k$ quan $k \in \mathbf{R}$

Proposició 2.3 *Sigui ϕ una forma bilineal (o sesquilineal) amb matriu B en la base $e_1, \ldots, e_n$. Designem per B_r el menor format per les r primeres files i les r primeres columnes de B. Aleshores ϕ és un producte escalar si i només si B és simètrica (o hermítica) i* $\det B_r > 0$ $\forall r$.

DEMOSTRACIÓ: Designem per E_r el subespai $\langle e_1, \ldots, e_r \rangle$ i per ϕ_r la restricció de ϕ a $E_r \times E_r$:

$$\begin{array}{rcl} \phi_r : E_r \times E_r & \longrightarrow & \mathbf{R} \quad (\text{o } \mathbf{C}) \\ (u, v) & \longmapsto & \phi(u, v). \end{array}$$

La matriu de ϕ_r en la base $e_1, \ldots, e_r$ és precisament B_r. Suposem que ϕ és un producte escalar a E. Llavors ϕ_r és un producte escalar a E_r i, per (2.1), existeix una base ortonormal de E_r. Si P és la matriu del canvi de la base $e_1, \ldots, e_r$ a la base ortonormal, tenim

$$I = P^t B_r \bar{P},$$

d'on $1 = \det B_r \cdot \mid \det P \mid^2$ i $\det B_r > 0$.

Suposem ara que $\det B_r > 0$, $\forall r$, i anem a construir una base $u_1, \ldots, u_n$ tal que $\phi(u_i, u_j) = 0$ si $i \neq j$, $\phi(u_i, u_i) = 1$. Aleshores (2.2) ens assegurarà que ϕ és un producte escalar.

Per construir la base "ortonormal" farem servir el mateix mètode de Gram-Schmidt utilitzat a (2.1).

- $\phi(e_1, e_1) = b_1^1 = \det B_1 > 0$. Per tant, existeix $\sqrt{\phi(e_1, e_1)}$ i podem construir

$$u_1 = \frac{e_1}{\sqrt{\phi(e_1, e_1)}},$$

que és un vector unitari, base de $E_1 = \langle e_1 \rangle$.

- Suposem que $u_1, \ldots, u_r$ és una base de E_r tal que

$$\begin{array}{rl} \phi(u_i, u_j) = 0 & \text{si } i \neq j \\ = 1 & \text{si } i = j. \end{array}$$

Com a (2.1) resulta que el vector

$$u'_{r+1} = e_{r+1} - (k^1 u_1 + \ldots + k^r u_r),$$

amb $k^i = \phi(e_{r+1}, u_i)$, és ortogonal a cada un dels vectors $u_1, \ldots, u_r$. Si demostrem que $\phi(u'_{r+1}, u'_{r+1}) > 0$, el vector unitari

$$u_{r+1} = \frac{u'_{r+1}}{\sqrt{\phi(u'_{r+1}, u'_{r+1})}}$$

serà tal que $u_1, \ldots, u_r, u_{r+1}$ formaran una base ortonormal de E_{r+1}. El resultat s'obté, llavors, per inducció. Calculem, doncs,

$$\begin{aligned} &\phi(u'_{r+1}, u'_{r+1}) = \\ &= \phi(e_{r+1}, e_{r+1}) - \phi(e_{r+1}, \sum_i k^i u_i) - \\ &\quad -\phi(\sum_i k^i u_i, e_{r+1}) + \phi(\sum_i k^i u_i, \sum_j k^j u_j) \end{aligned}$$

$$
\begin{aligned}
&= \phi(e_{r+1},e_{r+1}) - \sum_i \bar{k}^i\phi(e_{r+1},u_i) - \sum_i k^i\phi(u_i,e_{r+1}) + \sum_i k^i\bar{k}^i \\
&= \phi(e_{r+1},e_{r+1}) - \sum_i \overline{\phi(e_{r+1},u_i)}\phi(e_{r+1},u_i) \\
&\quad - \sum_i \phi(e_{r+1},u_i)\phi(u_i,e_{r+1}) + \sum_i \phi(e_{r+1},u_i)\overline{\phi(e_{r+1},u_i)} \\
&= \phi(e_{r+1},e_{r+1}) - \sum_i \phi(e_{r+1},u_i)\phi(u_i,e_{r+1}) \\
&= \begin{vmatrix} 1 & \cdots & 0 & \phi(u_1,e_{r+1}) \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & \phi(u_r,e_{r+1}) \\ \phi(e_{r+1},u_1) & \cdots & \phi(e_{r+1},u_r) & \phi(e_{r+1},e_{r+1}) \end{vmatrix}.
\end{aligned}
$$

(Això es pot comprovar, per exemple, desenvolupant per l'última fila).

La matriu que apareix aquí és la matriu de ϕ_{r+1} en la base $u_1,\ldots,u_r,e_{r+1}$ i s'obté, per tant, de la matriu B_{r+1} per un canvi de base. És, doncs, de la forma

$$P^tB_{r+1}\bar{P}$$

i el seu determinant és

$$\det B_{r+1} \mid \det P \mid^2 > 0$$

per ésser $\det B_{r+1} > 0$. Això acaba la demostració. □

Un *espai vectorial euclidià* és un espai vectorial sobre $\mathbf{R}$ amb un producte escalar. Un *espai vectorial unitari* és un espai vectorial sobre $\mathbf{C}$ amb un producte escalar. En tots dos casos és costum designar $\phi(u,v)$ per $\langle u,v\rangle$ o per $u\cdot v$.

Observació:

Sigui E un espai vectorial sobre $\mathbf{R}$ o $\mathbf{C}$ i sigui $u_1,\ldots,u_n$ una base qualsevol de E. Per (2.2) existeix sempre un producte escalar ϕ a E amb el qual $u_1,\ldots,u_n$ és una base ortonormal:

$$\phi(w,v) = w^1\bar{v}^1 + \ldots + w^n\bar{v}^n,$$

on $(w^1,\ldots,w^n)$, $(v^1,\ldots,v^n)$ són les coordenades de w, v en la base $u_1,\ldots,u_n$.

A $\mathbf{R}^n$ i $\mathbf{C}^n$ considerarem com a producte escalar estàndard aquell amb el qual la base $(1,0,\ldots,0),\ldots,(0,\ldots,0,1)$ és ortonormal.

XI.3 Norma

Sigui E un espai vectorial sobre $\mathbf{R}$ o $\mathbf{C}$. Una *norma* a E és una aplicació

$$\begin{array}{rccl} \| \ \| : & E & \longrightarrow & \mathbf{R} \quad \text{(sempre a } \mathbf{R}\text{!)} \\ & v & \longmapsto & \|v\| \end{array}$$

que compleix

1. $\|v\| = 0 \Leftrightarrow v = \vec{0}$;
2. $\|kv\| = |k| \cdot \|v\|$;
3. $\|u + v\| \leq \|u\| + \|v\|$. (Desigualtat triangular)

$|k|$ indica el valor absolut si $k \in \mathbf{R}$ i el mòdul si $k \in \mathbf{C}$.

Sigui E un espai vectorial amb un producte escalar ϕ. Farem servir la notació

$$\phi(u, v) = u \cdot v.$$

Sabem que $u \cdot u$ és sempre real positiu; designem per $\sqrt{u \cdot u}$ la determinació positiva de l'arrel quadrada de $u \cdot u$. L'aplicació $u \mapsto \sqrt{u \cdot u}$ és una norma a E. Per demostrar-ho necessitem un lema.

Lema 3.1 (Desigualtat de Cauchy-Schwarz) $\forall u, v \in E$,

$$| u \cdot v |^2 \leq (u \cdot u)(v \cdot v).$$

DEMOSTRACIÓ: Si $v = \vec{0}$, la desigualtat és certa. Suposem $v \neq \vec{0}$ i considerem

$$k = \frac{u \cdot v}{v \cdot v}.$$

Llavors

$$\begin{aligned} 0 &\leq (u - kv) \cdot (u - kv) = u \cdot u - k(v \cdot u) - \bar{k}(u \cdot v) + k\bar{k}(v \cdot v) \\ &= u \cdot u - \frac{(u \cdot v)(v \cdot u)}{(v \cdot v)} - \frac{\overline{(u \cdot v)}(u \cdot v)}{(v \cdot v)} + \frac{(u \cdot v)\overline{(u \cdot v)}}{(v \cdot v)} \\ &= u \cdot u - \frac{(u \cdot v)\overline{(u \cdot v)}}{(v \cdot v)} = u \cdot u - \frac{|u \cdot v|^2}{(v \cdot v)}, \end{aligned}$$

d'on $|u \cdot v|^2 \leq (u \cdot u)(v \cdot v)$.□

Proposició 3.2 *Sigui E un espai vectorial amb un producte escalar. L'aplicació*

$$\begin{array}{rccc} \| \ \| : & E & \longrightarrow & \mathbf{R} \\ & v & \longmapsto & \sqrt{v \cdot v} \end{array}$$

és una norma.

DEMOSTRACIÓ: La condició 1 de norma resulta del fet que el producte escalar és definit positiu. Per provar 2 observem que $(kv) \cdot (kv) = k\bar{k}(v \cdot v) = | k |^2 (v \cdot v)$.

Per demostrar 3 fem

$$\begin{aligned} (u + v) \cdot (u + v) &= u \cdot u + u \cdot v + v \cdot u + v \cdot v \\ &= u \cdot u + v \cdot v + (u \cdot v + \overline{u \cdot v}) \\ &\leq u \cdot u + v \cdot v + 2 \mid u \cdot v \mid \leq \text{ per (3.1)} \\ &\leq u \cdot u + v \cdot v + 2\sqrt{u \cdot u}\sqrt{v \cdot v} = (\sqrt{u \cdot u} + \sqrt{v \cdot v})^2, \end{aligned}$$

d'on $\sqrt{(u + v) \cdot (u + v)} \leq \sqrt{u \cdot u} + \sqrt{v \cdot v}$. □

XI.4 Producte escalar i espai dual

Sigui E un espai vectorial euclidià o unitari de dimensió finita. Per a tot $v \in E$ l'aplicació

$$\begin{array}{rcl} v^* : E & \longrightarrow & \mathbf{R} \text{ (o } \mathbf{C}) \\ u & \longmapsto & u \cdot v \end{array}$$

és lineal. Podem definir, doncs,

$$\begin{array}{rcl} \varphi : E & \longrightarrow & E' \\ v & \longmapsto & v^*, \end{array}$$

que compleix les següents propietats:

a) φ és injectiva, ja que $u^* = v^* \Rightarrow u^*(w) = v^*(w)\ \forall w \Rightarrow w \cdot u = w \cdot v\ \forall w \Rightarrow w \cdot (u - v) = 0\ \forall w \Rightarrow u - v = \vec{0} \Rightarrow u = v$.

b) φ és exhaustiva En efecte, donat $w \in E'$, considerem una base ortonormal $u_1, \ldots, u_n$ i el vector

$$u = \overline{w(u_1)}u_1 + \ldots + \overline{w(u_n)}u_n.$$

El vector u és una antiimatge de w, ja que $u^*(u_i) = u_i \cdot u = w(u_i)\ \forall i$; és a dir,

$$u^* = w.$$

c) $\varphi(v + u) = \varphi(v) + \varphi(u)$, ja que $(v + u)^*(w) = w(v + u) = w \cdot v + w \cdot u = v^*(w) + u^*(w) = (v^* + u^*)(w)\ \forall w$, i, per tant, $(v + u)^* = v^* + u^*$.

d) En el cas euclidià, $\varphi(kv) = k\varphi(v)$, ja que $(kv)^*(w) = w \cdot (kv) = k(w \cdot v) = k(v^*(w)) = (kv^*)(w)\ \forall w$, d'on $(kv)^* = kv^*$.

d') En el cas unitari, $\varphi(kv) = \bar{k}\varphi(v)$, ja que $(kv)^*(w) = w \cdot (kv) = \bar{k}(w \cdot v) = \bar{k}(v^*(w)) = (\bar{k}v^*)(w)\ \forall w$, d'on $(kv)^* = \bar{k}v^*$.

Hem demostrat, en particular, el

Teorema 4.1 *Si E és un espai euclidià, l'aplicació*

$$\begin{array}{rcl} \varphi : E & \longrightarrow & E' \\ v & \longmapsto & v^* \end{array}$$

és un isomorfisme canònic.□

Exercici:

Demostreu que si $u_1, \ldots, u_n$ és una base ortonormal, $u_1^*, \ldots, u_n^*$ és la seva base dual a E'.

XI.5 Subespais ortogonals

Sigui E un espai vectorial euclidià o unitari de dimensió finita i S un subconjunt de E. Anomenarem *subespai ortogonal de* S

$$S^{\perp} = \{v \in E \mid u \cdot v = 0 \; \forall u \in S\}.$$

Proposició 5.1 *Es compleix*

1. $S^{\perp}$ *és un subespai vectorial de* E;
2. $S \subset R \Rightarrow R^{\perp} \subset S^{\perp}$;
3. $S^{\perp} = \langle S \rangle^{\perp}$;
4. $\langle S \rangle \cap S^{\perp} = \{\vec{0}\}$;
5. $\langle S \rangle \subset (S^{\perp})^{\perp}$.

Exercici:

Demostreu (5.1).

Proposició 5.2 *Si* F *és un subespai vectorial de* E,

$$E = F \oplus F^{\perp}.$$

DEMOSTRACIÓ: De 4 de (5.1) es dedueix que $F \cap F^{\perp} = \{\vec{0}\}$.

Sigui $u_1, \ldots, u_n$ una base ortonormal de F. Completem-la fins a obtenir una base $u_1, \ldots, u_r, e_{r+1}, \ldots, e_n$ de E i apliquem el mètode de Gram-Schmidt per aconseguir una base $u_1, \ldots, u_r, u_{r+1}, \ldots, u_n$ ortonormal de E. Observem que $u_j \in F^{\perp}$ si $j = r+1, \ldots, n$. Aleshores, $\forall v \in E$,

$$v = (v^1 u_1 + \ldots + v^r u_r) + (v^{r+1} u_{r+1} + \ldots + v^n u_n) \in F + F^{\perp},$$

d'on resulta que $E = F \oplus F^{\perp}$.□

Corol·lari 5.3 $\dim F^{\perp} = \dim E - \dim F$. □

Corol·lari 5.4 *Si* F *és un subespai vectorial de* E, $F^{\perp\perp} = F$.

DEMOSTRACIÓ: Per 5. de (5.1), $F \subset F^{\perp\perp}$. Per (5.3), F i $F^{\perp\perp}$ tenen la mateixa dimensió. Per tant, $F = F^{\perp\perp}$.□

Observació:

La bijecció $\varphi : E \longrightarrow E'$ definida a l'apartat 4 aplica l'ortogonal d'un subespai F tal com l'acabem de definir, sobre l'ortogonal de F a E' definit a V.7. En efecte:

$$\{v \in E \mid u \cdot v = 0 \; \forall u \in F\} \xrightarrow{\varphi} \{v^* \in E' \mid v^*(u) = u \cdot v = 0 \; \forall u \in F\}.$$

Observeu que les propietats de $F^{\perp}$ demostrades a (5.1) i (5.2) són conseqüència immediata d'aquesta bijecció.

XI.6 Aplicacions adjuntes i autoadjuntes

Sigui E un espai vectorial euclidià o unitari. Un endomorfisme g es diu *l'aplicació adjunta* de $f \in \mathrm{End}(E)$ si

$$v \cdot g(u) = f(v) \cdot u \quad \forall u, v \in E.$$

L'adjunta, si existeix, és única. En efecte, si g_1 també és una adjunta de f tenim $v \cdot g(u) = f(v) \cdot u = v \cdot g_1(u)\ \forall v, u$, d'on $g(u) = g_1(u)\ \forall u$; per tant, $g = g_1$.

Existeix sempre l'adjunta de f? Anem a respondre a aquesta qüestió demostrant que l'adjunta de f no és res més que la dual $f' : E' \longrightarrow E'$ considerada com a aplicació de E a E via la bijecció φ de l'apartat 4. És a dir, veurem que l'aplicació

$$g = \varphi^{-1} f' \varphi : E \longrightarrow E' \longrightarrow E' \longrightarrow E$$

és lineal i $g(u) \cdot v = u \cdot f(v)\ \forall u, v$. En el cas real la linealitat de g és conseqüència de la linealitat de φ i de f'. En el cas complex la linealitat respecte a la suma de φ i f' implica la linealitat respecte a la suma de g, i si $k \in \mathbf{C}$

$$\begin{aligned} g(ku) = \quad \varphi^{-1} f' \varphi(ku) = \quad & \varphi^{-1} f'(\bar{k}\varphi(u)) = \quad \varphi^{-1}(\bar{k} f' \varphi(u)) \\ & = k(\varphi^{-1} f' \varphi(u)) = \quad kg(u). \end{aligned}$$

Observem ara quina és la imatge d'un vector $u \in E$ per g:

$$u \overset{\varphi}{\longmapsto} u^* \overset{f'}{\longmapsto} u^* \circ f \overset{\varphi}{\longleftarrow} g(u).$$

$g(u) \in E$ és tal que $g(u)^* = u^* \circ f$, d'on $g(u)^*(v) = u^* f(v)$, que equival a

$$v \cdot g(u) = f(v) \cdot u.$$

Proposició 6.1 *Si g és l'aplicació adjunta de $f \in \mathrm{End}(E)$ i $\dim E < \infty$,*

$$\mathrm{Nuc}\, g = (\mathrm{Im}\, f)^{\perp} \quad i \quad \mathrm{Im}\, g = (\mathrm{Nuc}\, f)^{\perp}.$$

DEMOSTRACIÓ: $\mathrm{Nuc}\, g = \{u \in E \mid g(u) = \vec{0}\} = \{u \in E; v \cdot g(u) = 0\ \forall v\}$ $= \{u \in E \mid f(v) \cdot u = 0\ \forall v\} = (\mathrm{Im}\, f)^{\perp}$. D'aquí, agafant ortogonals, obtenim $\mathrm{Im}\, f = (\mathrm{Nuc}\, g)^{\perp}$. Ara bé, si g és l'adjunta de f, f és l'adjunta de g i, en particular, es compleix $\mathrm{Im}\, g = (\mathrm{Nuc}\, f)^{\perp}$. □

Proposició 6.2 *Sigui $A = (a_i^j)$ la matriu de $f : E \longrightarrow E$ en una base ortonormal $u_1, \ldots, u_n$. La matriu de l'adjunta de f en la base $u_1, \ldots, u_n$ és A^t en el cas real i $\bar{A}^t$ en el cas complex.*

DEMOSTRACIÓ: Sigui g l'adjunta de f i $B = (b_i^j)$ la seva matriu. En el cas complex tenim

$$g(u_i) = \sum_{j=1}^{n} b_i^j u_j,$$

d'on

$$b_i^j = g(u_i) \cdot u_j = u_i \cdot f(u_j) = u_i \cdot \left(\sum_{k=1}^{n} a_j^k u_k \right) = \bar{a}_j^i.$$

Per tant, $B = \bar{A}^t$. El mateix raonament val en el cas real i s'obté $B = A^t$.□

Una *aplicació autoadjunta* $f \in \mathrm{End}(E)$ és una aplicació lineal que coincideix amb la seva adjunta; és a dir, tal que

$$v \cdot f(u) = f(v) \cdot u \qquad \forall u, v \in E.$$

De (6.2) resulta immediatament la

Proposició 6.3 *Si A és la matriu de f en una base ortonormal, f és autoadjunta si i només si*

$A = A^t$ *(A simètrica) en el cas real;*
$A = \bar{A}^t$ *(A hermítica) en el cas complex.*□

XI.7 Diagonalització de matrius simètriques i hermítiques

Tota matriu simètrica real o hermítica complexa és la matriu d'una aplicació autoadjunta en una base ortonormal. Això es dedueix de (6.3) i d'una observació feta al final de l'apartat 2. El problema de diagonalitzar aquestes matrius equival, doncs, al de trobar una base de vectors propis d'una aplicació autoadjunta.

Teorema 7.1 *Si E és un espai vectorial unitari i $f : E \longrightarrow E$ és autoadjunta, existeix una base de vectors propis ortonormal.*

DEMOSTRACIÓ: Procedirem per inducció sobre la dimensió de E. Si $\dim E = 1$, tot vector és propi i no hi ha res a demostrar. Si $\dim E = n$, el polinomi característic de f, $p(x) = det(f - xI) \in \mathbf{C}[x]$, té sempre una arrel. Sigui v un vector unitari de valor propi aquesta arrel λ: $f(v) = \lambda v$. El subespai

$$F = \langle v \rangle^{\perp} = \{u \in E \mid u \cdot v = 0\}$$

és invariant per f. En efecte, si $u \in F$,

$$f(u) \cdot v = u \cdot f(v) = v \cdot (\lambda v) = \bar{\lambda}(u \cdot v) = \bar{\lambda} \cdot 0 = 0,$$

d'on $f(u) \in F$. Per hipòtesi d'inducció existeix una base ortonormal i de vectors propis de F: $u_2, \ldots, u_n$. Aleshores $u_1 = v, u_2, \ldots, u_n$ és una base ortonormal de vectors propis de f (per (5.2)). □

Teorema 7.2 *Si E és un espai euclidià i $f : E \longrightarrow E$ és autoadjunta, existeix una base de vectors propis ortonormal.*

DEMOSTRACIÓ: La mateixa que en el cas unitari val sempre que puguem demostrar que el polinomi característic, $p(x) \in \mathbf{R}[x]$, té una arrel a $\mathbf{R}$. Això i més ens ho demostra el lema següent .□

Lema 7.3 *Sigui E un espai vectorial euclidià o unitari i sigui $f : E \longrightarrow E$ una aplicació autoadjunta. Aleshores el polinomi característic és de la forma*

$$p(x) = \pm(x - \lambda_1) \cdots (x - \lambda_n),$$

amb $\lambda_1, \ldots, \lambda_n \in \mathbf{R}$ (tant en el cas euclidià com en l'unitari).

DEMOSTRACIÓ: En el cas unitari la demostració es redueix a veure que tots els valors propis són reals. En efecte, si $f(v) = \lambda v$ amb $v \neq \vec{0}$, tenim $\lambda(v \cdot v) = (\lambda v) \cdot v = f(v) \cdot v = v \cdot f(v) = v \cdot (\lambda v) = \bar{\lambda}(v \cdot v)$. Com que $v \cdot v \neq 0$, cal que $\lambda = \bar{\lambda}$. És a dir, $\lambda \in \mathbf{R}$.

En el cas euclidià farem la demostració mitjançant una "complexificació" del problema. Sigui A la matriu simètrica corresponent a f en una certa base ortonormal. Si considerem A com una matriu complexa, A és hermítica i, per tant, és la matriu d'una certa aplicació lineal $\bar{f} : \bar{E} \longrightarrow \bar{E}$ d'un espai vectorial unitari. Com que f i $\bar{f}$ tenen la mateixa matriu, el seu polinomi característic serà el mateix i, per la primera part de la demostració, tindrà totes les arrels reals, com volíem demostrar.□

XI.8 Producte vectorial

Sigui E un espai vectorial euclidià de dimensió 3 i e_1, e_2, e_3 una base ortonormal de E. Fixats $u, v \in E$, l'aplicació

$$\begin{array}{ccc} E & \longrightarrow & \mathbf{R} \\ w & \longmapsto & \det_{(e_i)}(u, v, w) \end{array}$$

és lineal i, per tant, un element del dual E'. Sigui x el vector corresponent a aquest element en l'isomorfisme de (4.1)

$$\begin{array}{ccc} E & \longrightarrow & E' \\ x & \longmapsto & x^* = \det_{(e_i)}(u, v, \quad); \end{array}$$

és a dir, $\forall w$

$$w \cdot x = \det_{(e_i)}(u, v, w).$$

El vector x és, per definició, el *producte vectorial* de u i v, i el denotarem per

$$u \wedge v.$$

Proposició 8.1 *El producte vectorial compleix*

1. $w \cdot (u \wedge v) = \det_{(e_i)}(u, v, w)$;
2. $u \wedge v = -v \wedge u$;
3. $(ku) \wedge v = k(u \wedge v)$;
4. $(u + u') \wedge v = u \wedge v + u' \wedge v$;
5. $u \wedge v$ *és ortogonal a* u *i a* v;
6. $u \wedge v = \vec{0}$ *si i només si* u, v *són linealment dependents*;
7. *Si* $u \wedge v \neq \vec{0}$, u, v, $u \wedge v$ *és una base de la mateixa orientació que* e_1, e_2, e_3.

DEMOSTRACIÓ: 1 és la mateixa definició. Per demostrar 2 observem que $\forall w$, $w \cdot (u \wedge v) = \det_{(e_i)}(u, v, w) = \det_{(e_i)}(v, u, -w) = (-w) \cdot (v \wedge u) = w \cdot (-v \wedge u)$; d'on resulta que $u \wedge v = -v \wedge u$.

De manera anàloga es demostren 3 i 4; 5 resulta de 1:

$$u \cdot (u \wedge v) = \det_{(e_i)}(u, v, u) = 0$$

i, de la mateixa manera, $v \cdot (u \wedge v) = 0$. Si $u \wedge v = \vec{0}$, $\forall w$ $\det_{(e_i)}(u, v, w) = 0$. Si u i v fossin linealment independents, existiria un w tal que u, v, w seria una base i $\det_{(e_i)}(u, v, w) \neq 0$. Per tant, u i v són linealment dependents. El recíproc és clar. Així tenim 6. Per últim, si $u \wedge v \neq \vec{0}$, per 1,

$$\det_{(e_i)}(u, v, u \wedge v) = (u \wedge v) \cdot (u \wedge v) = \|u \wedge v\|^2 > 0,$$

que demostra 7.□

Proposició 8.2 *a)* $e_1 \wedge e_2 = e_3$, $e_2 \wedge e_3 = e_1$, $e_3 \wedge e_1 = e_2$.

b) Si $u = u^1 e_1 + u^2 e_2 + u^3 e_3$ *i* $v = v^1 e_1 + v^2 e_2 + v^3 e_3$,

$$u \wedge v = (u^2 v^3 - u^3 v^2) e_1 + (u^3 v^1 - v^3 u^1) e_2 + (u^1 v^2 - u^2 v^1) e_3.$$

Aquest resultat justifica la notació

$$u \wedge v = \begin{vmatrix} e_1 & u^1 & v^1 \\ e_2 & u^2 & v^2 \\ e_3 & u^3 & v^3 \end{vmatrix}$$

(desenvolupeu formalment per la primera columna!).

DEMOSTRACIÓ: De (8.1.1) resulta fàcilment que $e_i \cdot (e_1 \wedge e_2) = e_i \cdot e_3$ $\forall i$. Per tant, $e_1 \wedge e_2 = e_3$. Anàlogament es demostra $e_2 \wedge e_3 = e_1$ i $e_3 \wedge e_1 = e_2$. Llavors, per (8.1),

$$\begin{aligned} u \wedge v &= \textstyle\sum_{i,j} u^i v^j e_i \wedge e_j \\ &= (u^2 v^3 - u^3 v^2) e_1 + (u^3 v^1 - u^1 v^3) e_2 + (u^1 v^2 - u^2 v^1) e_3. \square \end{aligned}$$

Proposició 8.3 $(u \wedge v) \wedge w = (u \cdot w)v - (v \cdot w)u$.

DEMOSTRACIÓ: Si $u \wedge v = 0$, $u = kv$, d'on $(u \cdot w)v - (v \cdot w)u = k(v \cdot w)v - k(v \cdot w)v = 0$ i la igualtat de l'enunciat és certa. Si $u \wedge v \neq \vec{0}$, $u \wedge v \in \langle u, v \rangle^{\perp}$, d'on $(u \wedge v) \wedge w \in \langle u, v \rangle$. Per tant, $(u \wedge v) \wedge w = kv - hu$. Aquest vector ha d'esser també ortogonal a w, d'on resulta que

$$0 = (kv - hu) \cdot w; \quad \text{és a dir,} \quad k(v \cdot w) = h(u \cdot w).$$

Si $v \cdot w = 0$ i $u \cdot w = 0$, $w \in \langle u \wedge v \rangle$, d'on $(u \wedge v) \wedge w = \vec{0}$ i la igualtat de l'enunciat és certa. Si un dels dos productes és $\neq 0$, $k = a(u \cdot w)$ i $h = a(v \cdot w)$ per a un cert nombre real a. És a dir,

$$(u \wedge v) \wedge w = a((u \cdot w)v - (v \cdot w)u).$$

Per veure que $a = 1$, calculem la primera coordenada d'aquest vector:

$$\begin{aligned} &(u^3v^1 - u^1v^3)w^3 - (u^1v^2 - u^2v^1)w^2 \\ &\quad = a\big[(u^1w^1 + u^2w^2 + u^3w^3)v^1 - (v^1w^1 + v^2w^2 + v^3w^3)u^1\big]. \end{aligned}$$

D'aquí resulta que $a = 1$.□

El producte vectorial no és associatiu (com es pot deduir tant de 8.2 com de 8.3); en comptes de l'associativitat compleix la següent propietat:

Proposició 8.4 (Identitat de Jacobi)

$$(u \wedge v) \wedge w + (v \wedge w) \wedge u + (w \wedge u) \wedge v = \vec{0}.$$

DEMOSTRACIÓ: És conseqüència de (8.3). □

Proposició 8.5 $(u_1 \wedge u_2) \cdot (v_1 \wedge v_2) = (u_1 \cdot v_1)(u_2 \cdot v_2) - (u_1 \cdot v_2)(u_2 \cdot v_1)$.

DEMOSTRACIÓ:

$$\begin{aligned} &(u_1 \wedge u_2) \cdot (v_1 \wedge v_2) = \\ &\quad = \det_{(e_i)}(v_1, v_2, u_1 \wedge u_2) = \det_{(e_i)}(u_1 \wedge u_2, v_1, v_2) = v_2 \cdot ((u_1 \wedge u_2) \wedge v_1) \\ &\quad = v_2 \cdot ((u_1 \cdot v_1)u_2 - (u_2 \cdot v_1)u_1) = (u_1 \cdot v_1)(u_2 \cdot v_2) - (u_1 \cdot v_2)(u_2 \cdot v_1). \square \end{aligned}$$

Corol·lari 8.6 $\|u \wedge v\|^2 = \|u\|^2\|v\|^2 - (u \cdot v)^2$. □

(Això ens dóna el terme que falta a (3.1) en el cas particular d'un espai euclidià de dimensió 3).

Vegeu l'observació 2 de XII.5 per a una altra expressió de $\|u \wedge v\|$.

Observacions:

1. Les propietats 5 i 7 de (8.1) i el corol·lari (8.6) determinen el producte vectorial $u \wedge v$ de dos vectors linealment independents.
2. El producte vectorial depèn de la base ortonormal e_1, e_2, e_3. Considerem una altra base ortonormal u_1, u_2, u_3;

$$\det_{(u_i)}(u, v, w) = \det_{(u_i)}(e_1, e_2, e_3) \cdot \det_{(e_i)}(u, v, w).$$

Si les dues bases són de la mateixa orientació, $\det_{(u_i)}(e_1, e_2, e_3) = +1$ i els productes vectorials definits utilitzant una base i l'altra coincideixen. Si les bases són d'orientació oposada, $\det_{(u_i)}(e_1, e_2, e_3) = -1$ i els productes vectorials definits utilitzant una base i l'altra són vectors oposats.

XI.9 Nota històrica

Els primers passos en la teoria d'operadors lineals i en la introducció de productes escalars es deuen a Erhard Schmidt (1876–1959) el 1907, encara que es troben antecedents en els treballs de David Hilbert sobre les equacions integrals. Schmidt va introduir el concepte de norma, de producte escalar i d'ortogonalitat i va demostrar una generalització del teorema de Pitàgores i el fet que vectors ortogonals dos a dos són linealment independents.

Schmidt va desenvolupar la teoria utilitzant mètodes introduïts per Hermann Amandus Schwarz (1843–1921), el qual va demostrar, en particular, la desigualtat que porta el seu nom.

XI.10 Exercicis

1. Demostreu que si $\|u\| = \|v\|$ aleshores $(u+v) \cdot (u-v) = 0$. ($u+v$ i $u-v$ ens donen les bisectrius de u i v).

2. Demostreu la llei del paral·lelogram

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

3. Sigui G la matriu d'un producte escalar en una base e_1, e_2, e_3. Demostreu que $\det G = \det(e_1, e_2, e_3)^2$.

4. Sigui E l'espai vectorial dels polinomis reals de grau≤ 2. Per a tot parell $p(x), q(x) \in E$ definim

$$\phi(p, q) = \int_0^1 p(x) \cdot q(x)\, dx.$$

a) Demostreu que ϕ és un producte escalar a E.

b) Trobeu una base ortonormal de E.

c) Trobeu una base del subespai ortogonal al polinomi $2x + 1$.

5. Calculeu $\det(u \wedge v, v \wedge w, w \wedge u)$ en funció de $\det(u, v, w)$ (els determinants referits a una mateixa base).

6. Donats els dos vectors u, v de l'espai vectorial euclidià ordinari, considerem l'endomorfisme definit per

$$f(x) = (u \wedge x) \wedge v.$$

Comproveu que

a) Si $u = \vec{0}$ o $v = \vec{0}$, aleshores $f = 0$;

b) Si $u \neq \vec{0}$, $v \neq \vec{0}$ i $u \cdot v = 0$, aleshores $\operatorname{Nuc} f = \langle v \rangle^{\perp}$, $\operatorname{Im} f = \langle u \rangle$ i $f^2 = 0$;

c) Si $u \cdot v \neq 0$, aleshores $\operatorname{Nuc} f = \langle u \rangle$ i $\operatorname{Im} f = \langle v \rangle^{\perp}$;

d) Si $u \cdot v = 1$, f és una projecció.

7. Donats dos vectors u, v linealment independents de l'espai vectorial euclidià ordinari, definim un endomorfisme f per

$$f(x) = u \wedge (v \wedge x) + v \wedge (u \wedge x).$$

a) Demostreu que f és lineal i que $\det f = 2(u \cdot v)\|u \wedge v\|^2$.

b) Determineu els valors i vectors propis de f.

8. Siguin $a_1, \ldots, a_k$ punts donats de l'espai afí $\mathbf{R}^n$, $F = \langle \vec{a_1a_2}, \ldots, \vec{a_1a_k} \rangle$ i $u_{k+1}, \ldots, u_n$ una base de $F^{\perp}$. Si $U \in M_{n \times (n-k)}(\mathbf{R})$ té per columnes els vectors $u_{k+1}, \ldots, u_n$, demostreu que l'equació cartesiana de la varietat lineal determinada per $a_1, \ldots, a_k$ és $U^t x + b = \vec{0}$ on $b = -U^t a_1$.

9. Sigui E un espai vectorial euclidià, $f \in \operatorname{End}(E)$ tal que $\|f(x)\| \leq \|x\| \; \forall x \in E$ i g la seva adjunta.

Demostreu que

a) $\|g(x)\| \leq \|x\| \; \forall x \in E$;

b) $\operatorname{Nuc}(g - I) = \operatorname{Nuc}(f - I)$;

c) $E = \operatorname{Nuc}(f - I) + \operatorname{Im}(f - I)$.

10. Sigui E un espai vectorial euclidià i $f \in \operatorname{End}(E)$ tal que $f(x) \cdot y = -x \cdot f(y)$ $\forall x, y \in E$. Demostreu que

a) $\operatorname{Nuc} f$ i $\operatorname{Im} f$ són subespais ortogonals de E;

b) $E = \operatorname{Nuc} f \oplus \operatorname{Im} f$;

c) Si (a^i_j) és la matriu de f en una base ortonormal, aleshores $a^i_j = -a^j_i$ $\forall i, j$.

11. Sigui $f : E \longrightarrow F$ una aplicació lineal entre espais vectorials amb producte escalar.

Es diu *adjunta de* f una aplicació lineal $g : F \longrightarrow E$ tal que $v \cdot g(u) = f(v) \cdot u$ $\forall u \in F$, $v \in E$. Demostreu que g existeix i és única.

12. Sigui $f : E \longrightarrow F$ una aplicació lineal entre espais vectorials euclidians i g la seva adjunta (exercici 11). Demostreu que

a) $g \circ f$ és diagonalitzable en una base ortonormal;

b) Tots els valors propis de $g \circ f$ són positius. Designem-los per $a_1, \ldots, a_n$, $n = \dim E$;

c) Existeixen bases ortonormals $e_1, \ldots, e_n$ de E i $u_1, \ldots, u_n$ de F tals que $f(e_i) = \sqrt{a_i} u_i$, $i = 1, \ldots, n$.

XI.11 Exercicis de programar

13. Feu un programa que donada una matriu simètrica estudiï si és o no és definida positiva.

14. Sigui $e_1, \ldots, e_n$ una base donada de $\mathbf{R}^n$ i ϕ un producte escalar donat. Suposem que tenim les components dels e_i i la matriu de ϕ en la base canònica de $\mathbf{R}^n$. Feu un programa que permeti construir una base ortonormal seguint el mètode de Gram-Schmidt.

15. Feu un programa que, donada una família de vectors $e_1, \ldots, e_k$ de $\mathbf{R}^n$ i un producte escalar ϕ, trobi una base del subespai ortogonal a $F = \langle e_1, \ldots, e_k \rangle$. (per a la seva posterior utilització prepareu aquest programa com a subrutina).

(Indicació: hi ha dos mètodes

a) Resoldre el sistema homogeni

$$\begin{cases} (e_1)^t B x &= 0 \\ &\vdots \\ (e_k)^t B x &= 0, \end{cases}$$

on B és la matriu de ϕ;

b) Seguint la demostració de 5.2, extreure una base de F, completar-la a una base de $\mathbf{R}^n$ (exercici IV.18) i ortonormalitzar (exercici 14)).

16. Feu un programa que

a) Doni les equacions cartesianes $Ax + b = \vec{0}$, en la referència canònica, d'una varietat lineal de $\mathbf{R}^n$ a partir de l'expressió vectorial $a + F$;

b) Recíprocament, a partir de les equacions cartesianes trobi un punt a i una base de F.

(Indicació: Per a a), utilitzeu l'exercici 8. Per a b), el punt a és una solució del sistema $Ax + b = \vec{0}$ i una base de F és una base de solucions del sistema homogeni $Ax = \vec{0}$ (exercici VII.12)).

17. Feu un programa que donats $a_1, \ldots, a_k$ punts de $\mathbf{R}^n$ trobi les equacions cartesianes de la varietat lineal que generen.

(Indicació: feu servir els exercicis 8 i 16).

18. Prepareu un programa que donats $u, v \in \mathbf{R}^3$ calculi $u \wedge v$. Comproveu en exemples concrets les identitats del § 8.

Capítol XII

Aplicacions ortogonals. Aplicacions unitàries

El capítol anterior ha estat dedicat a l'estudi d'una nova estructura algebraica: els espais vectorials amb un producte escalar. Ara, com hem fet sempre que hem introduït una estructura algebraica, volem estudiar les aplicacions que conserven aquesta estructura: les aplicacions ortogonals i unitàries.

XII.1 Definicions

Sigui E un espai vectorial amb un producte escalar (euclidià o unitari). Una aplicació $f : E \longrightarrow E$ es diu *ortogonal*, en el cas real, o *unitària*, en el cas complex, si $\forall u,\ v \in E$

$$f(u) \cdot f(v) = u \cdot v.$$

Proposició 1.1 *Tota aplicació que conservi el producte escalar és lineal.*

DEMOSTRACIÓ: Per provar la linealitat respecte a la suma, anem a veure que, per a tot u, v, $f(u+v) - f(u) - f(v)$ té norma zero i, per tant, és el vector $\vec{0}$. En efecte,

$$\begin{aligned}
& (f(u+v) - f(u) - f(v)) \cdot (f(u+v) - f(u) - f(v)) \\
= \ & f(u+v) \cdot f(u+v) - f(u+v) \cdot f(u) - f(u+v) \cdot f(v) - f(u) \cdot f(u+v) + \\
& + f(u) \cdot f(u) + f(u) \cdot f(v) - f(v) \cdot f(u+v) + f(v) \cdot f(u) + f(v) \cdot f(v) = \\
= \ & (u+v) \cdot (u+v) - (u+v) \cdot u - (u+v) \cdot v - u \cdot (u+v) + \\
& + u \cdot u + + u \cdot v - v \cdot (u+v) + v \cdot u + v \cdot v = \\
= \ & ((u+v) - u - v) \cdot ((u+v) - u - v) = \vec{0} \cdot \vec{0} = 0.
\end{aligned}$$

Anàlogament tenim

$$\begin{aligned}
& (kf(u) - f(ku)) \cdot (kf(u) - f(ku)) = \\
& = k\overline{k} f(u) \cdot f(u) - kf(u) \cdot f(ku) - \overline{k} f(ku) \cdot f(u) + f(ku) \cdot f(ku) = \\
& = k\overline{k} u \cdot u - ku \cdot (ku) - \overline{k}(ku) \cdot u + (ku) \cdot (ku) = \\
& = (ku - ku) \cdot (ku - ku) = 0,
\end{aligned}$$

d'on $kf(u) - f(ku) = \vec{0}$ i $f(ku) = kf(u)$. □

Proposició 1.2 *Si f és ortogonal o unitària, es compleix*

1. $\|f(u)\| = \|u\| \ \forall u \in E$;
2. *v, u són ortogonals si i només si $f(v)$, $f(u)$ ho són;*
3. *f és bijectiva;*
4. *Si k és un valor propi de f, $|k| = 1$, on $|k|$ indica el valor absolut o el mòdul segons que sigui real o complex;*
5. *Si u, v són dos vectors propis de valors propis $k \neq h$, u, v són ortogonals.*

DEMOSTRACIÓ: 1 i 2 són conseqüència immediata de la conservació del producte escalar. 3 es dedueix de 1. Per demostrar 4, suposem $f(v) = kv$ amb $v \neq \vec{0}$. Aleshores

$$v \cdot v = f(v) \cdot f(v) = k\overline{k}(v \cdot v) = |k|^2(v \cdot v),$$

d'on $|k| = 1$.

Per provar 5, fem $u \cdot v = f(u) \cdot f(v) = k\overline{h}(u \cdot v)$. Si $u \cdot v \neq 0$, $k\overline{h} = 1$ i, com que $|h| = 1$, $\overline{h} = h^{-1}$. Per tant, $k = h$, en contra de la hipòtesi. Així doncs, $u \cdot v = 0$ i u i v són ortogonals. □

Proposició 1.3 *Sigui A la matriu de $f \in \mathrm{End}(E)$ en la base $e_1, \ldots, e_n$ de E. Aleshores*

$$f \text{ és } \left.\begin{matrix} \textit{ortogonal} \\ \textit{unitària} \end{matrix}\right\} \Leftrightarrow (a) \quad f(e_i) \cdot f(e_j) = e_i \cdot e_j \ \forall i, j$$

$$\Leftrightarrow (b) \quad \begin{cases} A^t G A = G \\ A^t G \overline{A} = G. \end{cases}$$

G indica la matriu del producte escalar en la base $e_1, \ldots, e_n$.

DEMOSTRACIÓ: de (a). $\Rightarrow$) és conseqüència de la conservació del producte escalar. $\Leftarrow$). Calculem $f(u) \cdot f(v)$. Si $u = \sum u^i e_i$ i $v = \sum v^i e_i$,

$$f(u) = \sum u^i f(e_i) \text{ i } f(v) = \sum v^i f(e_i),$$

i, per tant,

$$f(u) \cdot f(v) = \sum_{i,j} u^i \overline{v}^j f(e_i) \cdot f(ej) = \sum_{i,j} u^i \overline{v}^j e_i \cdot e_j = u \cdot v.$$

Anàlogament en el cas real.

Per demostrar (b) observem que $f(e_1), \ldots, f(e_n)$ és base per (1.2.(3)) i que la matriu del producte escalar en aquesta base és $A^t G A$ en el cas real i $A^t G \overline{A}$ en el cas complex. Si $f(e_i) \cdot f(e_j) = e_i \cdot e_j \ \forall i, j$, aquesta matriu coincideix amb la matriu del producte escalar en la base $e_1, ..., e_n$, que és G. □

Corol·lari 1.4 *Sigui A la matriu de $f \in \mathrm{End}(E)$ en una base ortonormal. Llavors,*
f és ortogonal $\Leftrightarrow$ $A^t A = I$;
f és unitària $\Leftrightarrow$ $A^t \overline{A} = I$. □

Corol·lari 1.5 *Sigui A la matriu del canvi d'una base ortonormal a una altra també ortonormal. Aleshores $A^t A = I$ (cas real) i $A^t \overline{A} = I$ (cas complex).* □

Una matriu real A es diu *ortogonal* si $A^t A = I$ (és a dir, si $A^{-1} = A^t$); llavors $\det A = \pm 1$. Una matriu complexa A es diu *unitària* si $A^t \overline{A} = I$ (és a dir, si $A^{-1} = \overline{A^t}$); llavors $|\det A| = 1$.

Les aplicacions ortogonals d'un espai vectorial euclidià E de dimensió n formen un grup que denominarem *grup ortogonal d'ordre n* i denotarem per $O(n)$. Les matrius reals $n \times n$ ortogonals formen un grup que, per (1.4), és isomorf al grup de les aplicacions ortogonals.

Les aplicacions unitàries d'un espai unitari E de dimensió n formen un grup que denominaren *grup unitari d'ordre n* i denotarem per $U(n)$. Les matrius complexes $n \times n$ unitàries formen un grup que, per (1.4), és isomorf al grup de les aplicacions unitàries.

Exercici:

Estudieu els grups $O(1)$ i $U(1)$.

XII.2 Diagonalització de matrius unitàries

Teorema 2.1 *Sigui E un espai vectorial unitari i sigui $f \in \mathrm{End}(E)$ una aplicació unitària. Existeix, sempre, una base ortonormal de E formada per vectors propis de f.*

DEMOSTRACIÓ: Per (VIII.2.5) sabem que existeix una base $v_1, \ldots, v_n$ en la qual f té una matriu triangular; és a dir, tal que

$$f(v_i) \in \langle v_1, \ldots, v_i \rangle.$$

Sigui $w_1, \ldots, w_n$ una base ortonormal obtinguda a partir de l'anterior pel mètode de Gram-Schmidt (XI.2.1). Com que $w_i \in \langle v_1, \ldots, v_i \rangle$, resulta que

$$f(w_i) \in \langle f(v_1), \ldots, f(v_i) \rangle = \langle v_1, \ldots, v_i \rangle = \langle w_1, \ldots, w_i \rangle$$

i, per tant, la matriu de f en la base $w_1, \ldots, w_n$ també és triangular.

Anem a veure que $w_1, \ldots, w_n$ són vectors propis. w_1 és, clarament, un vector propi. Suposem que $w_1, \ldots, w_k$ són vectors propis i sigui

$$f(w_{k+1}) = a^1 w_1 + \ldots + a^k w_k + a^{k+1} w_{k+1}.$$

Llavors tenim que, si $f(w_i) = b^i w_i$ per a $i = 1, \ldots, k$,

$$0 = w_{k+1} \cdot w_i = f(w_{k+1}) \cdot f(w_i) = f(w_{k+1}) \cdot (b^i w_i) = a^i \overline{b}^i.$$

Però $|b^i| = 1$ i, per tant, $b^i \neq 0$, d'on $a^i = 0$ per a $i = 1, ..., k$, i

$$f(w_{k+1}) = a^{k+1} w_{k+1}.$$

w_{k+1} també és, doncs, un vector propi. □

XII.3 Forma canònica d'una matriu ortogonal

Teorema 3.1 *Sigui E un espai vectorial euclidià i $f : E \longrightarrow E$ una aplicació ortogonal. Existeix una base ortonormal de E tal que la matriu corresponent a f és de la forma*

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & 0 & & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & A_1 & \\ & & 0 & & & & \ddots \\ & & & & & & & A_r \end{pmatrix},$$

on les matrius A_i són del tipus

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \text{ amb } a^2 + b^2 = 1.$$

DEMOSTRACIÓ: Siguin E_1 i E_{-1} els subespais de vectors propis de valor propi 1 i -1 respectivament. Els vectors de E_1 i E_{-1} són ortogonals entre ells (1.2.(5)) i, per tant, $E_1 \cap E_{-1} = \{\vec{0}\}$. Designem per F el subespai ortogonal a $E_1 \oplus E_{-1}$. És fàcil veure que F és invariant per f. La base buscada de

$$E = E_1 \oplus E_{-1} \oplus F$$

és unió de bases ortonormals de cada un d'aquests subespais. Com que F no conté vectors propis, el polinomi característic de la restricció de f a F és producte de polinomis irreductibles de segon grau. En particular, la dimensió de F és parella. Designem per A la matriu de la restricció de f a F en una certa base ortonormal $e_1, \ldots, e_{2r}$.

Sigui $\widehat{F}$ un espai vectorial unitari i $\widehat{e}_1, \ldots, \widehat{e}_{2r}$ una base ortonormal de $\widehat{F}$. Considerem l'aplicació

$$\widehat{f} : \widehat{F} \longrightarrow \widehat{F} \quad .$$

que en la base $\widehat{e}_1, \ldots, \widehat{e}_{2r}$ té la matriu A. Per (1.4) $\widehat{f}$ és unitària i per (2.1) existeix una base ortonormal de vectors propis. El polinomi característic de $\widehat{f}$ és el mateix

que el de f i les seves arrels són, per tant, conjugades dos a dos. Sigui v un vector propi unitari de valor propi $z = a + bi$. Si les coordenades de v en la base $\widehat{e}_1, \ldots, \widehat{e}_{2r}$ són $(x^1 + iy^1, \ldots, x^n + iy^n)$, designarem per $\overline{v}$ el vector de coordenades $(x^1 - iy^1, \ldots, x^n - iy^n)$. Aleshores, $\overline{v}$ és un vector unitari de valor propi $\overline{z}$. En efecte, amb la notació usual, tenim

$$A\overline{v} = \overline{Av} = \overline{zv} = \overline{z} \cdot \overline{v}.$$

Podem escollir, doncs, una base ortonormal $v_1, \ldots, v_{2r}$ de vectors propis de $\widehat{f}$ formada per parelles de vectors $v, \overline{v}$ amb valors propis conjugats, $z, \overline{z}$.

Considerem ara, per a cada parell $v, \overline{v}$, vectors

$$w = \tfrac{1}{\sqrt{2}}(v + \overline{v})$$

$$w' = \tfrac{-i}{\sqrt{2}}(v - \overline{v}).$$

Observem que w i w' tenen coordenades reals:

$$\begin{aligned} w &= \sqrt{2}(x^1, \ldots, x^n) \\ w' &= \sqrt{2}(y^1, \ldots, y^n). \end{aligned}$$

Tenint en compte que $v \cdot \overline{v} = 0$ (per 1.2 (5)) i que $v \cdot v = \overline{v} \cdot \overline{v} = 1$, resulta que

$$w \cdot w' = 0 \ \text{ i } \ \|w\| = 1 = \|w'\|.$$

A més a més,

$$\begin{aligned} \widehat{f}(w) &= \tfrac{1}{\sqrt{2}}\widehat{f}(v + \overline{v}) = \tfrac{1}{\sqrt{2}}(zv + \overline{zv}) = \\ &= \tfrac{1}{\sqrt{2}}\left((a + bi)\tfrac{1}{\sqrt{2}}(w + iw') + (a - bi)\tfrac{1}{\sqrt{2}}(w - iw')\right) = \\ &= aw - bw'. \end{aligned}$$

Anàlogament, $\widehat{f}(w') = bw + aw'$. És a dir, en llenguatge de matrius,

$$\begin{aligned} Aw &= aw - bw' \\ Aw' &= bw + aw'. \end{aligned}$$

Substituint en la base $v_1, \ldots, v_{2r}$ cada parella $v, \overline{v}$ pels corresponents w, w', obtenim una nova base ortonormal de $\widehat{F}$, $w_1, \ldots, w_{2r}$, formada per vectors de coordenades reals (en la base $\widehat{e}_1, \ldots, \widehat{e}_{2r}$).

Designem per u_i el vector de F que en la base $e_1, \ldots, e_{2r}$ té les mateixes coordenades que w_i en la base $\widehat{e}_1, \ldots, \widehat{e}_{2r}$. Clarament $u_1, \ldots, u_{2r}$ és una base ortonormal de F formada per parelles u, u' tals que

$$Au = au - bu'$$

$$Au' = bu + au'.$$

La matriu de f en aquesta base és de la forma

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ & & \ddots \\ 0 & & A_r \end{pmatrix},$$

on les caixes A_i són del tipus

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

amb $a^2 + b^2 = 1$. □

XII.4 Els grups $O(2)$ i $SO(2)$

En aquest apartat, E és un espai vectorial euclidià de dimensió 2. El conjunt d'aplicacions ortogonals de E amb la composicio és $O(2)$: grup ortogonal d'ordre 2. Per (1.4), si $f \in O(2)$, $\det f = \pm 1$. Considerem el morfisme

$$\begin{array}{ccc} O(2) & \longrightarrow & \{+1, -1\} \\ f & \longmapsto & \det f. \end{array}$$

El nucli d'aquest morfisme

$$SO(2) = \{f \in O(2) \mid \det f = +1\}$$

és un subgrup normal de $O(2)$ que es diu el *grup ortogonal especial* d'ordre 2. Començarem per estudiar aquest subgrup.

Fixem una base ortonormal e_1, e_2 de E. La matriu de $f \in SO(2)$

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

ha de complir $AA^t = I$ (o $A^{-1} = A^t$) (per (1.4)). És a dir,

$$A^{-1} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A^t$$

(ja que $detA = +1$) o sigui $a = d$, $c = -b$. Així, doncs,

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ amb } a^2 + b^2 = 1.$$

Proposició 4.1 *El grup $SO(2)$ és commutatiu.*

DEMOSTRACIÓ: Siguin $f, g \in SO(2)$ amb matrius en la base ortonormal e_1, e_2

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c & -d \\ d & c \end{pmatrix};$$

aleshores

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac-bd & -ad-bc \\ ad+bc & ac-bd \end{pmatrix} = \begin{pmatrix} c & -d \\ d & c \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

D'on $f \circ g = g \circ f$. □

Proposició 4.2 *Sigui $f \in SO(2)$ i*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

la seva matriu en la base ortonormal e_1, e_2. Aleshores, si u és un vector unitari qualsevol,

$$a = u \cdot f(u) \quad i \quad b = \det_{(e_i)}(u, f(u)).$$

DEMOSTRACIÓ: Sigui $u = \alpha e_1 + \beta e_2$ amb $\alpha^2 + \beta^2 = 1$. Llavors

$$f(u) = (a\alpha - b\beta)e_1 + (b\alpha + a\beta)e_2,$$

d'on

$$u \cdot f(u) = \alpha(a\alpha - b\beta) + \beta(b\alpha + a\beta) = a(\alpha^2 + \beta^2) = a,$$

i

$$\det_{(e_i)}(u, f(u)) = \begin{vmatrix} \alpha & a\alpha - b\beta \\ \beta & b\alpha + a\beta \end{vmatrix} = b(\alpha^2 + \beta^2) = b. \square$$

Aquesta proposició ens diu, en particular, que a és independent de la base ortonormal escollida i que b varia, com a molt, en el signe. En efecte, si u_1, u_2 és una altra base ortonormal,

$$\det_{(u_i)}(u, f(u)) = \det_{(u_i)}(e_1, e_2) \cdot \det_{(e_i)}(u, f(u)) = \pm b,$$

ja que $\det_{(u_i)}(e_1, e_2) = \pm 1$ (1.5). Aquest determinant és $+1$ quan e_1, e_2 i u_1, u_2 són de la mateixa orientació i -1 en cas contrari (IX.11).

Proposició 4.3 *Donats dos vectors $u, v \in E$ que tenen la mateixa norma, $\|u\| = \|v\| \neq 0$, existeix una $f \in SO(2)$ i només una tal que $f(u) = v$.*

DEMOSTRACIÓ: Sigui e_1, e_2 una base ortonormal. Si $f \in SO(2)$ compleix $f(u) = v$, la seva matriu en aquesta base ha d'ésser

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

on

$$a = \frac{u}{\|u\|} \cdot \frac{v}{\|v\|} \qquad b = \det_{e_i}\left(\frac{u}{\|u\|}, \frac{v}{\|v\|}\right).$$

Definim, doncs, f com l'aplicació lineal que té aquesta matriu. Per provar que f és ortogonal cal veure que $a^2 + b^2 = 1$. Sigui

$$u = \alpha e_1 + \beta e_2 \qquad v = \gamma e_1 + \delta e_2$$

Tenim

$$a^2 + b^2 = \frac{(u \cdot v)^2 + \det(u, v)^2}{\|u\|^2 \cdot \|v\|^2} = \frac{(\alpha\gamma + \beta\delta)^2 + (\alpha\delta - \gamma\beta)^2}{(\alpha^2 + \beta^2)(\gamma^2 + \delta^2)} = 1.$$

Demostrem ara que $f(u) = v$. Per (4.2),

$$\det_{(e_i)}\left(\frac{u}{\|u\|}, \frac{f(u)}{\|u\|}\right) = b = \det_{(e_i)}\left(\frac{u}{\|u\|}, \frac{v}{\|v\|}\right),$$

d'on $\det_{(e_i)}(u, f(u) - v) = 0$ i, per tant, $f(u) - v = k \cdot u$. Per altra banda,

$$\frac{u}{\|u\|} \cdot \frac{f(u)}{\|u\|} = a = \frac{u}{\|u\|} \cdot \frac{v}{\|v\|}$$

implica $u \cdot (f(u) - v) = 0$, d'on $u \cdot (ku) = 0$.

Per tant, $k = 0$ i $f(u) - v = 0$. □

Corol·lari 4.4 *Si $f \in SO(2)$ deixa un vector fix, aleshores $f = I$.* □

Estudiem ara les aplicacions $f \in O(2)$ amb $\det f = -1$.

Aquestes aplicacions formen una classe del quocient de $O(2)$ pel nucli $SO(2)$ de l'aplicació $\det : O(2) \longrightarrow \{+1, -1\}$ (III.5). Considerem, en particular, la que en la base ortonormal e_1, e_2 té matriu

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Qualsevol altra s'obté d'aquesta, component amb una aplicació de $SO(2)$, i la seva matriu serà, doncs, de la forma

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix} \qquad \text{amb } a^2 + b^2 = 1.$$

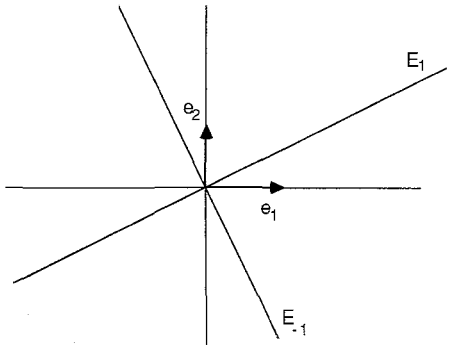
Observem que

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Es tracta, per tant, de simetries (X.2.III). Els subespais de vectors propis de valor propi $+1$ i -1 són, respectivament,

$$E_1 = \langle (b, 1 - a) \rangle \quad \text{i} \quad E_{-1} = \langle (-b, 1 + a) \rangle.$$

Aquests subespais són ortogonals i, per això, aquestes simetries es diuen *simetries ortogonals.*



XII.5 Angles

Sigui E un espai vectorial euclidià de dimensió 2. En el conjunt de parells de vectors unitaris definim una relació d'equivalència de la següent forma:

$$\begin{aligned}(u,u') \sim (v,v') &\Leftrightarrow \text{existeix } f \in SO(2) \text{ tal que } f(u)=v, f(u')=v'.\\ &\Leftrightarrow \text{existeix } g \in SO(2) \text{ tal que } g(u)=u', g(v)=v'.\end{aligned}$$

Demostrem primer que aquestes dues condicions són equivalents.

$\Rightarrow$). Suposem que existeix l'aplicació f i sigui $g \in SO(2)$ tal que $g(u)=u'$ (4.3). Aleshores, per la commutativitat de $SO(2)$,

$$g(v) = gf(u) = fg(u) = f(u') = v'.$$

$\Leftarrow$). Suposem ara que existeix g i sigui $f \in SO(2)$ tal que $f(u)=v$. Llavors

$$f(u') = fg(u) = gf(u) = g(v) = v'.$$

Anomenarem *angle* cada una de les classes d'equivalència per aquesta relació. L'angle determinat per un parell (u,u') el denotarem per $\widehat{uu'}$ o simplement per (u,u'). L'angle de dos vectors u,v no necessàriament unitaris, $\widehat{uv}$, és l'angle dels vectors $\frac{u}{\|u\|}, \frac{v}{\|v\|}$.

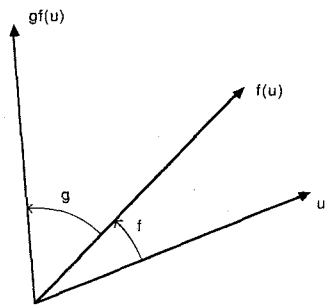
Designarem per A el conjunt d'angles. L'aplicació

$$\begin{array}{ccc} SO(2) & \longrightarrow & A \\ f & \longmapsto & \widehat{uf(u)} \end{array}$$

on u és qualsevol vector unitari, és una bijecció que ens permet transportar l'operació de $SO(2)$ a A: donats $\alpha,\beta \in A$ amb antiimatges f i g, respectivament, definim la *suma* $\alpha+\beta$ com l'angle corresponent a $g \circ f$.

Així, si u és un vector unitari,

$$\left.\begin{array}{rcl}\alpha &=& (u,f(u))\\ \beta &=& (f(u),gf(u))\end{array}\right\} \Rightarrow \alpha+\beta = (u,gf(u)).$$



Naturalment, la suma definida a A té les mateixes propietats que l'operació de $SO(2)$. A és, doncs, un grup commutatiu amb element neutre $0 = \widehat{uu}$. L'oposat d'un angle $\widehat{uf(u)}$ és $\widehat{f(u)u}$.

Fixada una orientació a E, a cada $f \in SO(2)$ li correspon per (4.2) una matriu ben determinada

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{amb } a^2 + b^2 = 1.$$

Sigui α l'angle corresponent a f. Definim el *cosinus de* α ($\cos\alpha$) i el *sinus de* α ($\sin\alpha$) com els valors a i b en aquesta matriu:

$$\cos\alpha = a \qquad \sin\alpha = b.$$

Observem que en canviar l'orientació de l'espai E canvia el signe de $\sin\alpha$ però no el de $\cos\alpha$. A més, es té

$$\cos^2\alpha + \sin^2\alpha = 1.$$

L'angle 0 correspon a l'aplicació identitat; per tant,

$$\cos 0 = 1 \qquad \sin 0 = 0.$$

L'angle oposat de α correspon a l'aplicació inversa f^{-1}, que té per matriu la transposada de la matriu de f; per tant,

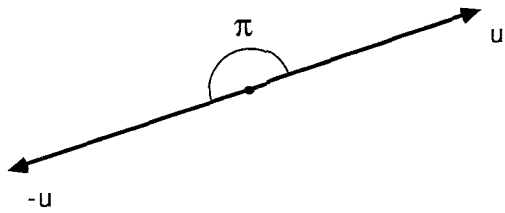
$$\cos(-\alpha) = \cos\alpha \qquad \sin(-\alpha) = -sin(\alpha).$$

L'angle $\alpha + \beta$ correspon a la composició de les aplicacions corresponents a α i β; s'obté

$$\begin{aligned} \cos(\alpha+\beta) &= \cos\alpha \cdot \cos\beta - \sin\alpha \cdot \sin\beta \\ \sin(\alpha+\beta) &= \sin\alpha \cdot \cos\beta + \cos\alpha \cdot \sin\beta \end{aligned}$$

Proposició 5.1 *Existeix un angle $\pi \neq 0$, i només un, tal que $2\pi = 0$. π és l'angle tal que*

$$\cos\pi = -1 \qquad i \qquad \sin\pi = 0.$$



DEMOSTRACIÓ: Sigui $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = A$ la matriu corresponent a π en una orientació prefixada. La condició $2\pi = 0$ equival a $AA = I$; és a dir, $A^{-1} = A$. Però, com que A és ortogonal ($A^{-1} = A^t$), $A = A^t$. D'on resulta que $b = 0$ i $a = \pm 1$. En el cas a=1, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ correspon a l'angle 0.

En el cas $a = -1$, $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ correspon a l'angle buscat. □

Observem que $\pi = \widehat{u(-u)}$, on u és un vector qualsevol. π es diu *l'angle pla.*

Proposició 5.2 *Existeixen dos angles δ_1, δ_2, i només dos, tals que $2\delta_i = \pi$. δ_1 i δ_2 són els angles amb*

$$\cos\delta_1 = cos\delta_2 = 0 \qquad \sin\delta_1 = -\sin\delta_2 = 1.$$

DEMOSTRACIÓ: Sigui

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

la matriu de l'angle δ buscat. La condició $2\delta = \pi$ equival a $AA = -I$; és a dir, $A^{-1} = -A$. Per ser A ortogonal, $A^{-1} = A^t$, d'on $A^t = -A$ i, per tant, $a = 0$, $b = \pm 1$. □

Els angles δ_1, δ_2 es diuen els *angles rectes.*

Exercici:

$\widehat{uv}$ és un angle recte si i només si u i v són ortogonals.

Proposició 5.3 *Donat un angle α, existeixen dos angles ϕ_1, ϕ_2, i només dos, tals que $2\phi = \alpha$. A més a més,*

$$\cos\alpha = \cos^2\phi - \sin^2\phi \qquad i \qquad \sin\alpha = 2\cos\phi \cdot \sin\phi.$$

DEMOSTRACIÓ: Siguin

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \qquad B = \begin{pmatrix} c & -d \\ d & c \end{pmatrix}$$

les matrius de α i ϕ respectivament. La condició $2\phi = \alpha$ equival a $BB = A$:

$$\begin{pmatrix} c^2 - d^2 & -cd \\ 2cd & c^2 - d^2 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

De $c^2 - d^2 = a$, $c^2 + d^2 = 1$ resulta que $c = \pm\sqrt{\frac{1+a}{2}}$, $d = \pm\sqrt{\frac{1-a}{2}}$, d'on

$$2cd = \pm\sqrt{1 - a^2} = \pm\sqrt{b^2} = \pm|b|.$$

Però $2cd = b$. Per tant, si $b > 0$, c i d han d'ésser tots dos positius o tots dos negatius; si $b < 0$, c i d han d'ésser un positiu i l'altre negatiu. En ambdós casos hi ha dues solucions, com es tractava de veure. El cas $b = 0$ ja ha estat considerat a (5.1) i (5.2). □

Observacions:

1. Sigui $\alpha = \widehat{uv}$. Si u i v no són unitaris, recordem que $\widehat{uv}$ és, per definició, l'angle de $\frac{u}{\|u\|}$ i $\frac{v}{\|v\|}$. Per (4.2) i la definició de $\cos\alpha$ tenim que

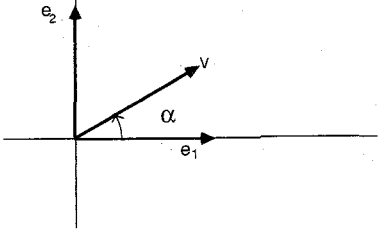
$$\cos\alpha = \frac{u}{\|u\|} \cdot \frac{v}{\|v\|}.$$

   D'on $u \cdot v = \|u\| \cdot \|v\| \cdot \cos\alpha$.

2. De l'observació anterior i de (XI.8.6) resulta que

$$\|u \wedge v\| = \|u\| \cdot \|v\| \cdot |\sin\widehat{uv}|.$$

3. Sigui e_1, e_2 una base ortonormal i $v = ae_1 + be_2$ un vector unitari ($a^2 + b^2 = 1$).



   L'angle $\alpha = \widehat{e_1 v}$ correspon a una aplicació $f \in SO(2)$ tal que $f(e_1) = v$. La seva matriu serà

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Per tant, $a = \cos\alpha$ i $b = \sin\alpha$. És a dir,

$$v = \cos\alpha \cdot e_1 + \sin\alpha \cdot e_2.$$

L'aplicació $f \in SO(2)$ corresponent a un angle α es diu *rotació (vectorial) d'angle* α. $SO(2)$ és el *grup de les rotacions* de E. Observem que la traça de qualsevol matriu de f és $2\cos\alpha$. Així, α és l'angle tal que

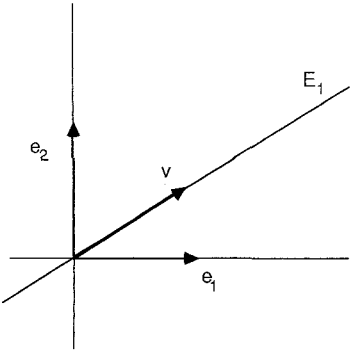
$$\cos\alpha = \frac{1}{2}\ \mathrm{tr}\ f.$$

El signe de $\sin\alpha$ no queda determinat; depèn de l'orientació de la base en la qual treballem.

Sigui ara $f \in O(2)$ amb $\det f = -1$ i matriu, en una base ortonormal e_1, e_2,

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \qquad a^2 + b^2 = 1.$$

f és una simetria ortogonal d'eix $E_1 = \langle (b, 1-a) \rangle$.



Suposem $a \neq 1$. Considerem el vector unitari

$$v = \left(\frac{b}{\sqrt{2(1-a)}}, \frac{1-a}{\sqrt{2(1-a)}} \right).$$

Observem que la seva segona coordenada és positiva ($a < 1$). Posem $\alpha = \widehat{e_1 v}$. Aleshores

$$\left.\begin{array}{l} \cos\alpha = \frac{b}{\sqrt{2(1-a)}} \\ \sin\alpha = \frac{1-a}{\sqrt{2(1-a)}} \end{array}\right\} \Rightarrow \begin{array}{lcll} \cos 2\alpha & = & \cos^2\alpha - \sin^2\alpha & = a \\ \sin 2\alpha & = & 2\sin\alpha \cdot \cos\alpha & = b \end{array}.$$

La matriu de f és, per tant, de la forma

$$\begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}.$$

Si $a = 1$, llavors $v = e_1$, $\alpha = 0$, i val el mateix resultat.

Donat un angle $\alpha = \widehat{uv}$ i una aplicació $f \in \mathrm{End}(E)$ designarem per $f\alpha$ l'angle $\widehat{f(u)f(v)}$.

Si $f \in SO(2)$, i $g \in SO(2)$ és l'aplicació corresponent a l'angle α, tenim

$$f\alpha = \widehat{f(u)fg(u)} = \widehat{f(u)gf(u)} = \alpha.$$

Si f és una simetria ortogonal és molt fàcil veure que

$$fg = g^{-1}f.$$

Aleshores

$$f\alpha = \widehat{f(u)fg(u)} = \widehat{f(u)g^{-1}f(u)} = -\alpha.$$

La proposició següent resumeix aquests dos fets.

Proposició 5.4 *Les rotacions vectorials conserven els angles. Les simetries ortogonals els inverteixen.* □

XII.6 El grup $O(3)$

En aquest apartat E és un espai vectorial euclidià de dimensió 3. Com en el cas de dimensió 2 (§4),

$$SO(3) = \{f \in O(3) \mid \det f = +1\}.$$

Estudiem primer les aplicacions de $SO(3)$. El polinomi característic de $f \in SO(3)$ és de grau 3 i té, per tant, una arrel real. Sigui $v \neq \vec{0}$ un vector propi. El subespai $F = \langle v \rangle^{\perp}$ és invariant per f i la restricció de f a F és ortogonal. Si v és de valor propi -1, el determinant d'aquesta restricció és -1; seria, doncs, una simetria ortogonal i, en particular, tindria el valor propi $+1$. Resulta, per tant, que $+1$ és sempre valor propi de f.

Sigui u_3 un vector propi unitari de valor propi $+1$ i sigui u_1, u_2, u_3 una base ortonormal. La restricció

$$f : \langle u_1, u_2 \rangle = \langle u_3 \rangle^{\perp} \longrightarrow \langle u_3 \rangle^{\perp}$$

té determinant $+1$ i és una rotació vectorial (d'angle α). La matriu de f en la base u_1, u_2, u_3 és, aleshores,

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Direm llavors que f és una *rotació vectorial d'eix* $\langle u_3 \rangle$ *i angle* α. Observem que $\cos\alpha$ queda determinat per la traça:

$$\cos\alpha = \frac{1}{2}(\operatorname{tr} f - 1).$$

El signe de $\sin\alpha$ depèn de l'orientació de la base u_1, u_2 de $\langle u_3 \rangle^{\perp}$.

Atenció!:

El signe de $\sin\alpha$ no depèn de l'orientació de u_1, u_2, u_3.

L'eix de f és al nucli de l'aplicació

$$\phi = f - f^{-1}.$$

Sigui $A = (a_i^j)$ la matriu de f en una base ortonormal e_1, e_2, e_3. La matriu de f^{-1} és $A^{-1} = A^t$ i la de ϕ

$$A - A^t = \begin{pmatrix} 0 & a_2^1 - a_1^2 & a_3^1 - a_1^3 \\ a_1^2 - a_2^1 & 0 & a_3^2 - a_2^3 \\ a_1^3 - a_3^1 & a_2^3 - a_3^2 & 0 \end{pmatrix}.$$

- Si $\phi \neq 0$, $\operatorname{Nuc}\phi = \langle (a_2^3 - a_3^2), (a_3^1 - a_1^3), (a_1^2 - a_2^1) \rangle$ és l'eix de f.
- Si $\phi = 0$ (és a dir, si $A = A^t$ és simètrica), llavors $f = f^{-1}$ i $f^2 = I$. Les úniques aplicacions de $SO(3)$ que compleixen això són la identitat i una rotació d'angle π o *simetria axial*. Per a tot v,

$$f(v + f(v)) = f(v) + f^2(v) = f(v) + v$$

és de l'eix. En particular,

$$e_1 + f(e_1) = (a_1^1 + 1, a_1^2, a_1^3)$$
$$e_2 + f(e_2) = (a_2^1, a_2^2 + 1, a_2^3)$$
$$e_3 + f(e_3) = (a_3^1, a_3^2, a_3^3 + 1)$$

són de l'eix. Observem que aquests tres vectors no poden ser simultàniament $\vec{0}$, ja que llavors $f = -I \notin SO(3)$.

Observació:

És fàcil veure que una aplicació lineal $\Phi : E \longrightarrow E$ té una matriu de la forma

$$\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

en una base ortonormal si i només si , per a tot $v \in E$, $v \cdot \phi(v) = 0$. Existeix, aleshores, un vector $u \in E$ ($u = (-c, b, -a) \in \operatorname{Nuc}\Phi$) tal que

$$\Phi = u \wedge -;$$

és a dir,

$$\Phi(v) = u \wedge v \qquad \forall v \in E.$$

Passem ara a estudiar el conjunt de $f \in O(3)$ amb $\det f = -1$. El mateix raonament fet per demostrar que, si $f \in SO(3)$, f té el valor propi $+1$, prova que si $f \in O(3)$ amb $\det f = -1$, f té el valor propi -1. Sigui u_3 un vector propi unitari de valor propi -1 i sigui u_1, u_2, u_3 una base ortonormal. La restricció

$$f : \langle u_1, u_2 \rangle = \langle u_3 \rangle^{\perp} \longrightarrow \langle u_3 \rangle^{\perp}$$

és una aplicació ortogonal amb determinant $+1$; és a dir, una rotació a $\langle u_3 \rangle^{\perp}$. La matriu de f en la base u_1, u_2, u_3 és de la forma

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

En particular, si $\alpha = 0$ direm que es tracta d'una *simetria especular* (vectorial) respecte al subespai $\langle u_3 \rangle^{\perp}$. En general, f és composició d'una simetria especular i una rotació d'eix ortogonal al pla de simetria.

La matriu de f que acabem de trobar ens diu que

$$\cos\alpha = \frac{1}{2}(\ \mathrm{tr}\ f + 1).$$

Tant el pla de simetria $\langle u_3 \rangle^{\perp}$ com l'eix de rotació estan determinats pel subespai de vectors propis de valor propi -1. Aquest subespai està contingut en el nucli de

$$\phi = (f - f^{-1}).$$

- Si $\phi \neq 0$, obtenim, com en el cas del grup ortogonal especial,

$$\mathrm{Nuc}\,\phi = \langle (a_2^3 - a_3^2), (a_3^1 - a_1^3), (a_1^2 - a_2^1) \rangle = \langle u_3 \rangle.$$

- Si $\phi = 0$ (A simètrica: $A = A^t = A^{-1}$), $f^2 = I$. Les úniques simetries ortogonals amb $\det f = -1$ són les simetries especulars i la *simetria central* $f = -I$. Per a tot v,

$$f(f(v) - v) = f^2(v) - f(v) = -(f(v) - v).$$

En particular,

$$f(e_1) - e_1 = (a_1^1 - 1, a_1^2, a_1^3)$$

$$f(e_2) - e_2 = (a_2^1, a_2^2 - 1, a_2^3)$$

$$f(e_3) - e_3 = (a_3^1, a_3^2, a_3^3 - 1)$$

són ortogonals al pla de simetria.

XII.7 Una altra determinació de les rotacions

Una rotació $f \in SO(3)$ queda determinada per l'eix $\langle u \rangle$ i l'angle α. Aquest, per la seva banda, està determinat per una orientació de $\langle u \rangle^\perp$ i per $\sin\alpha$ i $\cos\alpha$.

Suposem fixada una orientació a l'espai E. Aleshores, tota orientació de $\langle u \rangle$ n'indueix una a $\langle u \rangle^\perp$ i viceversa, de la següent manera: si $v \neq 0$ té orientació positiva a $\langle u \rangle$, una base $v_1, v_2 \in \langle u \rangle^\perp$ és d'orientació positiva a $\langle u \rangle^\perp$ si i només si v_1, v_2, v és d'orientació positiva a E.

Considerem l'eix orientat de forma que $\sin\alpha \geq 0$. Sigui $v \in \langle u \rangle$ amb

$$\|v\| = \left|\sin\frac{\alpha}{2}\right|$$

i orientació positiva. v ens dóna, per tant, l'orientació de $\langle u \rangle^\perp$ i el valor de $\cos\alpha$ i $\sin\alpha$:

$$\begin{aligned} \cos\alpha &= \cos^2\frac{\alpha}{2} - \sin^2\frac{\alpha}{2} = 1 - 2\|v\|^2 \\ \sin\alpha &= 2\cos\tfrac{\alpha}{2}\cdot\sin\tfrac{\alpha}{2} = 2\|v\|\sqrt{1-\|v\|^2}. \end{aligned}$$

La rotació queda, doncs, completament determinada per v i la denotarem per g_v. Observem que qualsevol vector v amb $\|v\| \leq 1$ determina una rotació.

Quan $\sin\alpha = 0$, qualsevol de les dues orientacions de l'eix compleix $\sin\alpha \geq 0$. Aquest cas es presenta si $\alpha = 0$ o $\alpha = \pi$.

- $\alpha = 0 \Rightarrow \|v\| = 0 \Rightarrow v = \vec{0}$.
- $\alpha = \pi \Rightarrow \|v\| = 1$ i v pot ser qualsevol dels dos vectors unitaris de $\langle u \rangle$. Llavors $g_v = g_{-v}$ és una simetria axial que denotarem per s_v.

XII.8 Composició de rotacions

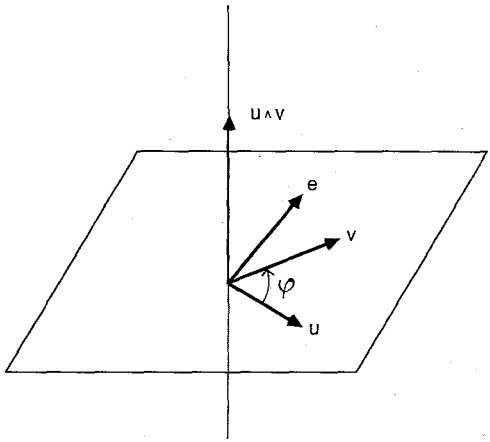
Si dues rotacions tenen el mateix eix és molt fàcil veure que la seva composició és una rotació amb aquest eix i angle la suma dels angles de les dues rotacions. La determinació dels elements que caracteritzen la rotació és molt més complexa en el cas de rotacions amb eixos diferents. L'objectiu d'aquest apartat és, donades dues rotacions g_{v_1}, g_{v_2} (amb la notació del §7), calcular el vector w tal que $g_w = g_{v_2} \circ g_{v_1}$.

Començarem per estudiar la composició de dues simetries axials.

Proposició 8.1

$$\begin{aligned} s_v \circ s_u &= g_{u\wedge v} \quad si \quad u\cdot v \geq 0 \\ &= g_{v\wedge u} \quad si \quad u\cdot v \leq 0. \end{aligned}$$

DEMOSTRACIÓ: $u \wedge v$ ortogonal a u i a $v \Rightarrow s_v \circ s_u(u \wedge v) = s_v(-u \wedge v) = u \wedge v \Rightarrow \langle u \wedge v \rangle$ és de l'eix de $s_v \circ s_u$. $u \wedge v = \vec{0} \Leftrightarrow u = \pm v$(ja que tots dos són unitaris) $\Leftrightarrow s_u = s_v \Leftrightarrow s_v \circ s_u = I = g_{\vec{0}}$.



Suposem $u \wedge v \neq \vec{0}$. L'angle de la rotació $s_v \circ s_u$ és l'angle que formen un vector qualsevol de $\langle u, v\rangle$, per exemple u, i la seva imatge: $s_v s_u(u) = s_v(u)$. Aquest angle és precisament $\alpha = 2\widehat{uv}$; en efecte, la matriu de s_v en una base ortonormal $\{u, e\}$ és, pel §5,

$$\begin{pmatrix} \cos 2\varphi & \sin 2\varphi \\ \sin 2\varphi & -\cos 2\varphi \end{pmatrix},$$

on φ és l'angle de v amb l'eix: $\varphi = \widehat{uv}$. Per tant,

$$s_v(u) = \cos 2\varphi \cdot u + \sin 2\varphi \cdot e,$$

d'on $\alpha = \widehat{us_v(u)} = 2\varphi$.

Per altra banda, $\|u \wedge v\| = |\sin \widehat{uv}| = |\sin \frac{\alpha}{2}|$, d'on resulta que el vector que determina la rotació $s_v \circ s_u$ és $u \wedge v$ o $v \wedge u$.

Observem que amb la notació de més amunt

$$\det_{(u,e)}(u, v) = \sin \varphi$$

i, per tant, $\sin \varphi$ és positiu o negatiu segons que l'orientació de $\langle u, v\rangle$ sigui la donada per u, v o l'oposada. Aleshores tenim que

- si $u \cdot v > 0 \Leftrightarrow \cos \varphi > 0 \Rightarrow \sin \alpha = 2 \sin \varphi \cos \varphi \geq 0$ només si $\sin \varphi \geq 0 \Rightarrow$ l'orientació a $\langle u, v\rangle$ ha d'ésser la donada per $u, v \Rightarrow$ l'orientació de l'eix ha d'ésser la de $u \wedge v \Rightarrow$

$$s_v \circ s_u = g_{u \wedge v}.$$

- si $u \cdot v < 0 \Leftrightarrow \cos\varphi < 0 \Rightarrow \sin\alpha \geq 0$ només si $\sin\varphi \leq 0 \Rightarrow$ l'orientació de $\langle u, v\rangle$ ha d'ésser la de $v, u \Rightarrow$ l'orientació de l'eix és la de $v \wedge u \Rightarrow$

$$s_v \circ s_u = g_{v\wedge u}.$$

- si $u \cdot v = 0$, llavors $\|u \wedge v\| = 1$ i per tant $g_{u\wedge v} = g_{v\wedge u}$ torna a ser una simetria axial. □

La proposició (8.1) suggereix que tota rotació g_v es pot descompondre en producte de dues simetries axials d'eixos perpendiculars a l'eix de la rotació. De fet una de les simetries pot escollir-se arbitràriament.

Proposició 8.2 *Donada una simetria axial* s_e *(*$\|e\| = 1$*) i una rotació* g_v *amb* $v \cdot e = 0$, *existeixen simetries axials* s_w *i* s_u *tals que*

$$g_v = s_e \circ s_w \qquad g_v = s_u \circ s_e.$$

DEMOSTRACIÓ: Si $v = \vec{0} \Rightarrow g_v = I \Rightarrow s_w = s_e,\ s_u = s_e$.

Suposem $v \neq \vec{0}$. Per (8.1) hem de buscar $w \in \langle v\rangle^{\perp} = \langle e, v \wedge e\rangle$. Escrivim

$$w = ae + b(v \wedge e).$$

Com que $s_w = s_{-w}$, podem suposar que $a = e \cdot w \geq 0$. Aleshores, per (8.1),

$$v = w \wedge e = b(v \wedge e) \wedge e = b((v \cdot e)e - (e \cdot e)v) = -bv.$$

D'on $b = -1$. A més a més,

$$1 = \|w\|^2 = a^2 + \|v \wedge e\|^2 = a^2 + \|v\|^2 \Rightarrow a = \sqrt{1 - \|v\|^2}.$$

Però $\|v\| = |\sin\frac{\alpha}{2}|$, on α és l'angle de la rotació g_v. Per tant, $a = |\cos\frac{\alpha}{2}|$. És a dir, si w existeix cal que:

$$w = |\cos\frac{\alpha}{2}|e - v \wedge e.$$

És immediat comprovar que efectivament aquest compleix $s_e \circ s_w = g_v$.

Per un raonament anàleg a aquest s'obté

$$u = |\cos\frac{\alpha}{2}|e + v \wedge e. \ \square$$

El problema de trobar la rotació composició de dues g_{v_1}, g_{v_2} amb eixos no paral·lels ($v_1 \wedge v_2 \neq \vec{0}$) es redueix, per (8.3), a un problema de composició de simetries axials que (8.2) resol. En efecte, si e és un vector unitari ortogonal a v_1 i a v_2, (8.3) ens dóna vectors u_1, u_2 tals que

$$g_{v_1} = s_e \circ s_{u_1} \qquad g_{v_2} = s_{u_2} \circ s_e.$$

Aleshores, per (8.2),

$$g_{v_2} \circ g_{v_1} = s_{u_2} \circ s_{u_1} \begin{array}{lll} = & g_{u_1 \wedge u_2} & \text{si} \quad u_1 \cdot u_2 \geq 0 \\ = & g_{u_2 \wedge u_1} & \text{si} \quad u_1 \cdot_2 \leq 0. \end{array}$$

Calculem aquests vectors. Per a això escollim un vector e concret $e = \dfrac{v_1 \wedge v_2}{\|v_1 \wedge v_2\|}$. Llavors, si α_1, α_2 són els angles de les rotacions g_{v_1}, g_{v_2},

$$u_1 = |\cos \frac{\alpha_1}{2}| e + e \wedge v_1 \qquad \text{i} \qquad u_2 = |\cos \frac{\alpha_2}{2}| e - e \wedge v_2,$$

d'on

$$\begin{aligned} u_1 \wedge u_2 &= |\cos \frac{\alpha_1}{2}| v_2 + |\cos \frac{\alpha_2}{2}| v_1 - (e \wedge v_1) \wedge (e \wedge v_2) = \\ &\quad |\cos \frac{\alpha_1}{2}| v_2 + |\cos \frac{\alpha_2}{2}| v_1 - (v_1 \wedge v_2) \\ u_1 \cdot u_2 &= |\cos \frac{\alpha_1}{2}| |\cos \frac{\alpha_2}{2}| - (e \wedge v_1)(e \wedge v_2) = \\ &= |\cos \frac{\alpha_1}{2}| |\cos \frac{\alpha_2}{2}| - v_1 \cdot v_2 = \\ &= |\cos \frac{\alpha_1}{2}| |\cos \frac{\alpha_2}{2}| - |\sin \frac{\alpha_1}{2}| |\sin \frac{\alpha_2}{2}| \cos \widehat{v_1 v_2} \end{aligned}$$

Aquestes expressions de $u_1 \wedge u_2$ i $u_1 \cdot u_2$ són vàlides també si les dues rotacions tenen el mateix eix: $v_1 \wedge v_2 = \vec{0}$. Llavors, en els càlculs anteriors, e és un vector unitari qualsevol ortogonal a l'eix i $(e \wedge v_1) \wedge (e \wedge v_2) = \vec{0} = v_1 \wedge v_2$. Tenim, doncs, el següent resultat:

Proposició 8.3 *Siguin g_{v_1}, g_{v_2} dues rotacions amb angles α_1 i α_2 respectivament. Si el valor de*

$$|\cos \frac{\alpha_1}{2}| |\cos \frac{\alpha_2}{2}| - |\sin \frac{\alpha_1}{2}| |\sin \frac{\alpha_2}{2}| \cos \widehat{v_1 v_2}$$

és positiu,

$$g_{v_2} \circ g_{v_2} = g_w,$$

on

$$w = |\cos \frac{\alpha_1}{2}| v_2 + |\cos \frac{\alpha_2}{2}| v_1 - (v_1 \wedge v_2).$$

Si aquell valor és negatiu,

$$g_{v_2} \circ g_{v_2} = g_{-w}$$

per al mateix w. □

XII.9 Nota històrica

Malgrat que el nom "matriu ortogonal" ja fou utilitzat el 1854 per Charles Hermite (1822–1901), no fou fins el 1878 que Georg Ferdinand Frobenius (1849–1917) en va donar una definició formal, demostrant-ne les primeres propietats. Hermite habia considerat ja el 1855 les matrius hermítiques, demostrant que els valors propis són reals, resultat demostrat per Arthur Buckheim (1859–1888) el 1885 per a les matrius simètriques.

XII.10 Exercicis

1. Demostreu que si F és un subespai vectorial de E invariant per $f \in O(E)$, $F^\perp$ també és invariant per f.

2. Demostreu que per a tota $A \in M_{n\times n}(\mathbf{R})$, $A = (a_i^j)$, $\sum_{i,j}(a_i^j)^2$ és invariant per canvis de base ortogonals.
   (Indicació: proveu que $\sum_{i,j}(a_i^j)^2 = \text{traça}(AA^t)$).

3. Estudieu el subgrup de $O(2)$ que deixa invariant el conjunt format per dos subespais vectorials de dimensió 1.

4. Estudieu els elements de $O(3)$ que deixen invariant el conjunt format per dos plans ortogonals.

5. Donats dos vectors $u \neq v$ d'igual norma d'un espai vectorial euclidià de dimensió 2, estudieu les aplicacions ortogonals f tals que $f(u) = v$ i $f(v) = u$.

6. En una base u, v una rotació vectorial té per matriu
$$\begin{pmatrix} 1 & -3/2 \\ 2/3 & 0 \end{pmatrix}.$$
   Determineu l'angle de la rotació, l'angle dels vectors u, v i la raó $\dfrac{\|v\|}{\|u\|}$.

7. Sigui f una simetria axial de $\mathbf{R}^2$ d'eix $\langle u\rangle$ i v un vector qualsevol. Demostreu que $\widehat{vf(v)} = 2\widehat{vu}$.

8. Sigui
$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$$
   una matriu simètrica i $D = Q^{-1}AQ$ la seva forma diagonal. Demostreu que Q és una rotació d'angle
$$\phi = \frac{1}{2}\arctan\frac{2b}{d-a}.$$

9. Sigui $f \in \mathrm{End}(E)$ tal que, $\forall u, v \in E$,

$$f(u) \cdot f(v) = \lambda(u \cdot v) \qquad (\lambda > 0 \text{ fix}).$$

Demostreu que:

a) $\|f(v)\| = \mu\|v\|$ amb $\mu = +\sqrt{\lambda}$, $\forall v \in E$;

b) f és bijectiva;

c) Els únics valors propis possibles de f són $\pm\mu$;

d) $f = g \circ h$ on $g \in O(n)$ i h és una homotècia.

10. Demostreu que si $f \in \mathrm{End}(\mathbf{R}^2)$ té en una base ortonormal una matriu de la forma

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

aleshores existeix un $\lambda \in \mathbf{R}$ tal que

$$f(u) \cdot f(v) = \lambda(u \cdot v) \quad \forall u, v \in \mathbf{R}^2.$$

Calculeu λ i determineu la descomposició de l'apartat d) de l'exercici anterior.

11. Demostreu que tota matriu quadrada real A admet una descomposició $A = QR$ amb Q ortogonal i R triangular superior.

(Indicació: demostreu primer que si $B \in M_{2\times 2}(\mathbf{R})$ existeix sempre una rotació P tal que PB és triangular superior. Llavors, donada $A \in M_{n\times n}(\mathbf{R})$, es poden anar eliminant tots els elements subdiagonals amb rotacions bidimensionals adients).

Demostreu que si $\det A \neq 0$, la descomposició $A = QR$ és única.

12. Demostreu que tot endomorfisme d'un espai vectorial euclidià de dimensió finita es pot descompondre en producte d'un autoadjunt i un ortogonal. Estudieu la unicitat de la descomposició.

(Indicació: feu servir l'exercici XI.12)

XII.11 Exercicis de programar

13. Feu un programa que donats dos vectors u, v de $\mathbf{R}^2$ calculi l'angle $\widehat{uv}$.

14. Utilitzant l'exercici XI.14 prepareu una llista de matrius de $O(3)$ per tal de fer-les servir d'exemples allà on convingui.

15. Prepareu un programa que permeti

a) Donada $A \in O(3)$, trobar l'eix i l'angle llevat del signe (§6).

b) Donada $A \in SO(3)$, trobar el vector $v \in \mathbf{R}^3$ tal que $A = g_v$ (§7).

c) Donat $v \in \mathbf{R}^3$ unitari, trobar la matriu de la rotació g_v en la base canònica.

16. Donades dues rotacions g_u, g_v de $\mathbf{R}^3$, calcular-ne la composició $g_u \circ g_v$ (8.3).

Proposem a continuació quatre mètodes iteratius per a calcular valors propis reals de matrius (veure l'exercici VIII.26).

17. Mètode QR

Escriviu $A = QR$ on Q és ortogonal i R és triangular superior (exercici 11). Sigui $A_1 = RQ$. L'algorisme consisteix a anar descomponent $A_k = Q_k R_k$ i fent $A_{k+1} = R_k Q_k$, $k = 1, 2, \ldots$ Les matrius A_k són totes equivalents a A i tendeixen cap a una matriu triangular superior, sota hipòtesis més generals que en el mètode LU.

18. Mètode de Jacobi (S'aplica només quan la matriu és simètrica).

Donada A simètrica, cada rotació bidimensional Q com la de l'exercici 11 anul·la un parell d'elements de A. La rotació següent no respecta els zeros aconseguits, però és fàcil deduir de l'exercici 2 que el valor $\sum_{i \neq j} (a_i^j)^2$ disminueix a cada rotació.

Prepareu el programa de la següent manera: sigui a_i^j l'element de A de mòdul màxim entre els que no estan a la diagonal principal. Posem $A_1 = {Q_1}^{-1} A Q_1$, on Q_1 és una rotació bidimensional que anul·la a_i^j. Repetint el procés s'obté una successió $A_k = {Q_k}^{-1} A Q_k$ que convergeix cap a una matriu diagonal.

Observem que el mètode ens dóna també una base ortonormal de vectors propis.

19. Mètode de la potència

Donada $A \in M_{n \times n}(\mathbf{R})$ diagonalitzable a $\mathbf{C}$ i tal que els seus valors propis satisfan $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$, aquest algorisme permet aproximar el valor propi dominant λ_1 i el vector propi u_1 corresponent.

Trieu un vector $x_0 \in \mathbf{R}^n$ qualsevol ($x_0 \neq \vec{0}$). Considereu la successió

$$x_{k+1} = A\left(\frac{x_k}{\|x_k\|}\right) \quad k = 0, 1, 2, \ldots$$

És senzill veure que $\{x_k\}$ convergeix cap al vector propi u_1. El valor propi λ_1 es pot obtenir a partir d'una component qualsevol no nul·la de x_k, ja que, per a k gran, Ax_k aproxima $\lambda_1 u_1$.

Aquest mètode també és vàlid si λ_1 té multiplicitat més gran que 1. En aquest cas $\{x_k\}$ convergeix cap a algun vector del subespai de vectors propis de valor propi λ_1 (depenent del x_0 inicial).

20. Mètode d'iteració inversa

Una senzilla variant del mètode anterior permet aproximar tots els valors propis reals de A i els vectors propis corresponents.

Només cal observar que si $\alpha \in \mathbf{R}$ és un paràmetre qualsevol, aleshores, si $\alpha \neq \lambda_i$,

$$Au_i = \lambda_i u_i \iff (A - \alpha I)^{-1} u_i = \frac{1}{\lambda_i - \alpha} u_i.$$

Llavors, si λ és el valor propi de A més proper a α, resulta que $1/(\lambda - \alpha)$ és valor propi dominant de $(A - \alpha I)^{-1}$. Per tant, la successió

$$\begin{cases} x_0 \neq \vec{0} & \text{qualsevol} \\ x_{k+1} = (A - \alpha I)^{-1} \left(\frac{x_k}{\|x_k\|} \right) & k = 0, 1, 2, \ldots \end{cases}$$

tendeix cap al vector propi u tal que $Au = \lambda u$. Aleshores, els quocients x_{k+1}^i / x_k^i tendeixen cap a $1/(\lambda - \alpha)$.

Així, donant valors a α, podem anar "caçant" un a un els valors propis reals de A.

Capítol XIII

Espais afins euclidians

En l'estudi dels espais afins fet al capítol IX no apareixen conceptes tan usuals en la geometria clàssica com distàncies, perpendicularitat i angles. El motiu és que aquestes nocions estan relacionades amb l'existència d'un producte escalar a l'espai vectorial associat a l'espai afí. En aquest capítol, estudiarem els espais afins reals que tenen associat un espai vectorial euclidià: els espais afins euclidians.

XIII.1 Espais afins euclidians

Un espai afí real (A, E) es diu un *espai afí euclidià* si a E hi ha un producte escalar, és a dir, si E és un espai vectorial euclidià.

Donats dos punts d'un espai afí euclidià, $p, q \in A$, es diu *distància* entre p i q el nombre real

$$\mathrm{d}(p,q) = \|\overrightarrow{pq}\|.$$

L'aplicació $\mathrm{d} : A \times A \longrightarrow \mathbf{R}$ que assigna a cada parell $(p,q) \in A \times A$ el nombre real $\mathrm{d}(p,q)$ es diu *aplicació distància* i compleix les següents propietats, $\forall p, q, r \in A$:

1. $\mathrm{d}(p,q) \geq 0$; $\mathrm{d}(p,q) = 0 \Leftrightarrow p = q$.
2. $\mathrm{d}(p,q) = \mathrm{d}(q,p)$.
3. $\mathrm{d}(p,q) \leq \mathrm{d}(p,r) + \mathrm{d}(r,q)$. (Desigualtat triangular).
4. $\mathrm{d}(p,q) \geq |\mathrm{d}(p,r) - \mathrm{d}(r,q)|$.

Les tres primeres propietats són conseqüència de propietats de la norma. Per demostrar la quarta observem que, per 3 i 2,

$$\begin{cases} \mathrm{d}(p,q) + \mathrm{d}(q,r) \geq \mathrm{d}(p,r) \Rightarrow \mathrm{d}(p,q) \geq \mathrm{d}(p,r) - \mathrm{d}(r,q) \\ \mathrm{d}(r,p) + \mathrm{d}(p,q) \geq \mathrm{d}(r,q) \Rightarrow \mathrm{d}(p,q) \geq \mathrm{d}(r,q) - \mathrm{d}(p,r). \end{cases}$$

Exercici:

Proveu que $\mathrm{d}(p,q) = \mathrm{d}(p,x) + \mathrm{d}(x,q)$ si i només si el punt x és del segment $\overline{pq}$.

Proposició 1.1 (Teorema de Pitàgores) *Siguin p,q,r tres punts d'un espai afí euclidià (A,E). Si $\overrightarrow{pq} \cdot \overrightarrow{pr} = 0$, es compleix*

$$d(p,q)^2 + d(p,r)^2 = d(q,r)^2.$$

DEMOSTRACIÓ: $\mathrm{d}(q,r)^2 = \|\overrightarrow{qr}\|^2 = \overrightarrow{qr} \cdot \overrightarrow{qr} = (\overrightarrow{pr} - \overrightarrow{pq}) \cdot (\overrightarrow{pr} - \overrightarrow{pq}) = \|\overrightarrow{pr}\|^2 + \|\overrightarrow{pq}\|^2 = \mathrm{d}(p,r)^2 + \mathrm{d}(p,q)^2$. □

Sigui (A,E) un espai afí euclidià de dimensió n.

Dues varietats, $a+F$ i $b+G$, tals que qualsevol vector $u \in F$ és ortogonal a qualsevol vector $v \in G$ ($u \cdot v = 0$), direm que són *ortogonals* o *perpendiculars*. Aleshores $F \subset G^\perp$ i, per tant, la suma de les dimensions d'aquestes varietats és $\leq n$. Dues varietats $a+F$ i $b+G$ tals que $\dim F + \dim G \geq n$ direm que són *ortogonals* o *perpendiculars* si les varietats $a+F^\perp$ i $b+G^\perp$ són ortogonals, és a dir, si $F^\perp \subset G$.

Exemples:

1. Sigui $a_1x^1 + \ldots + a_nx^n = b$ l'equació d'un hiperplà en un sistema de referència $\{p;\ e_1,\ldots,e_n\}$ ortonormal (tal que la base $e_1,\ldots,e_n$ és ortonormal). El vector de coordenades $(a_1,\ldots,a_n)$ és ortogonal a qualsevol vector $(x^1,\ldots,x^n)$ de la direcció de H, ja que aquests vectors compleixen $a_1x^1 + \ldots + a_nx^n = 0$.

2. Sigui A l'espai afí euclidià de dimensió 3. Dues rectes $a+\langle w\rangle$ i $b+\langle v\rangle$ es creuen si no són paral·leles ni es tallen. Llavors w,v són linealment independents i $\overrightarrow{ab} \notin \langle w,v\rangle$ (IX.4.1).

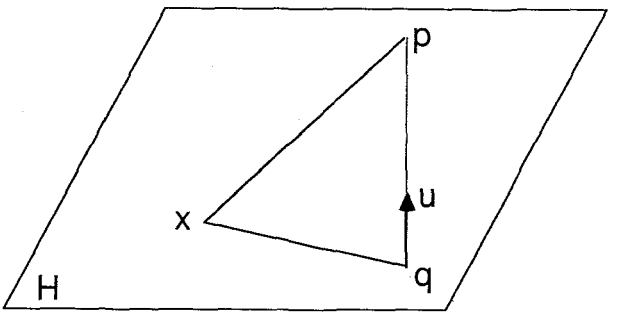
   Considerem els plans que contenen una de les rectes i la direcció perpendicular a ambdues. Si $\vec{0} \neq u \in (w,v)^\perp$, aquests plans són

   $$a + \langle w,u\rangle$$
   $$b + \langle v,u\rangle.$$

   Con que w,v,u formen una base, $\overrightarrow{ab} \in \langle w,u\rangle + \langle v,u\rangle$ i els plans es tallen en una recta $c+\langle u\rangle$. Aquesta recta talla $a+\langle w\rangle$, ja que $\overrightarrow{ac} \in \langle w,u\rangle$, i talla $b+\langle v\rangle$, ja que $\overrightarrow{bc} \in \langle v,u\rangle$. La recta $c+\langle u\rangle$ es diu la *perpendicular comuna* a $a+\langle v\rangle$ i $b+\langle v\rangle$; els punts d'intersecció amb aquestes rectes es diuen els *peus de la perpendicular comuna.*

3. Sigui $H = a+F$ un hiperplà i p un punt tal que $p \notin H$. Com que $F \oplus F^\perp = E$, la recta $p+F^\perp$ talla H exactament en un punt (IX, §4). Aquest punt s'anomena *projecció ortogonal de p sobre H.*

   Si en diem q, llavors $\forall x \in H$ es té $\overrightarrow{pq} \cdot \overrightarrow{qx} = 0$, ja que $\overrightarrow{pq} \in F^\perp$ i $\overrightarrow{qx} \in F$.



XIII.2 Distància entre dues varietats lineals

Donades dues varietats lineals L_1, L_2 el conjunt de les distàncies entre els seus punts té un mínim que denominarem *distància entre* L_1 *i* L_2:

$$\mathrm{d}(L_1, L_2) = \mathrm{mín}\{\mathrm{d}(p, q) \mid p \in L_1,\ q \in L_2\}.$$

L'existència d'aquest mínim és conseqüència de la continuïtat de l'aplicació distància i del fet que el conjunt de distàncies està acotat inferiorment per 0. A més a més, es pot veure també que sempre existeixen punts $a \in L_1$, $b \in L_2$ tals que $\mathrm{d}(a, b) = \mathrm{d}(L_1, L_2)$. Nosaltres no anem a demostrar aquí cap d'aquests fets en general; ens limitarem a un parell de casos particulars interessants. El tractament d'aquests casos és, per altra banda, un bon model per a trobar la distància entre altres tipus de varietats.

I Distància d'un punt p a un hiperplà $H = a + F$
Sigui q la projecció ortogonal de p sobre H. Aleshores, $\forall x \in H$,

$$\mathrm{d}(p, x)^2 = \mathrm{d}(p, q)^2 + \mathrm{d}(q, x)^2 \geq \mathrm{d}(p, q)^2 \quad \text{(per 1.1)},$$

ja que $\overrightarrow{pq} \cdot \overrightarrow{qx} = 0$. Tenim doncs que $\mathrm{d}(p, q)$ és el mínim de les distàncies del punt p als punts de H: $\mathrm{d}(p, H) = \mathrm{d}(p, q)$. Per trobar concretament el valor d'aquesta distància considerem un vector unitari u ortogonal a la direcció de H. Llavors $\overrightarrow{pq} = ku$ i $\mathrm{d}(p, q) = |k|$. Sigui $x \in H$ qualsevol;

$$\overrightarrow{px} \cdot u = (\overrightarrow{pq} + \overrightarrow{qx}) \cdot u = \overrightarrow{pq} \cdot u + \overrightarrow{qx} \cdot u = ku \cdot u = k,$$

ja que $\overrightarrow{qx} \cdot u = 0$. Per tant,

$$\mathrm{d}(p, H) = \left|\overrightarrow{px} \cdot u\right|,$$

on $x \in H$ és qualsevol i u és un vector unitari ortogonal a la direcció de H.

Si $a_1x^1 + \ldots + a_nx^n + b = 0$ és l'equació de H en un sistema de referència ortonormal, triem

$$u = \frac{1}{r}(a_1, \ldots, a_n),$$

on $r = \|(a_1, \ldots, a_n)\| = \sqrt{a_1{}^2 + \ldots + a_n{}^2}$. Així doncs,

$$\begin{aligned} \overrightarrow{px} \cdot u &= \frac{1}{r}\left((x^1 - p^1)a_1 + \ldots + (x^n - p^n)a_n\right) = \\ &= \frac{1}{r}\left(x^1a_1 + \ldots + x^na_n - p^1a_1 - \ldots - p^na_n\right) = \\ &= \frac{-1}{r}(b + p^1a_1 + \ldots + p^na_n), \end{aligned}$$

i, per tant,

$$\mathrm{d}(p, H) = \frac{\left|a_1p^1 + \ldots + a_np^n + b\right|}{\sqrt{a_1{}^2 + \ldots + a_n{}^2}}.$$

Observem que el numerador d'aquesta expressió és el valor de l'equació de H en el punt $p = (p^1, \ldots, p^n)$. En particular, si $p \in H$, llavors òbviament $\mathrm{d}(p, H) = 0$.

II Distància entre dues rectes que es creuen, en un espai afí euclidià de dimensió 3

Siguin $r = a + \langle w \rangle$ i $s = b + \langle v \rangle$ aquestes dues rectes, p, q els peus de la perpendicular comuna i $p + \langle u \rangle$ aquesta perpendicular (§1). Suposem u unitari. Considerem el pla $a + \langle w, v \rangle$ i la projecció ortogonal b' de b sobre aquest pla: b' és la intersecció de $b + \langle u \rangle$ amb $a + \langle w, v \rangle$. Aleshores $\overrightarrow{bb'}$ i $\overrightarrow{b'a}$ són ortogonals i $\mathrm{d}(a, b) \geq \mathrm{d}(b', b)$ per (1.1). Ara bé, el punt $c = b + \overrightarrow{qp}$ és de $b + \langle u \rangle$ i de $p + \langle w, v \rangle = a + \langle w, v \rangle$, ja que $\overrightarrow{pc} = \overrightarrow{pb} + \overrightarrow{qp} = \overrightarrow{qb} \in \langle v \rangle$. Per tant, $b' = c = b + \overrightarrow{qp}$, d'on $\overrightarrow{b'b} = \overrightarrow{qp}$ i $\mathrm{d}(b', b) = \mathrm{d}(p, q)$. Així doncs, $\mathrm{d}(p, q)$ és la mínima distància entre els punts de r i s:

$$\mathrm{d}(r, s) = \mathrm{d}(p, q).$$

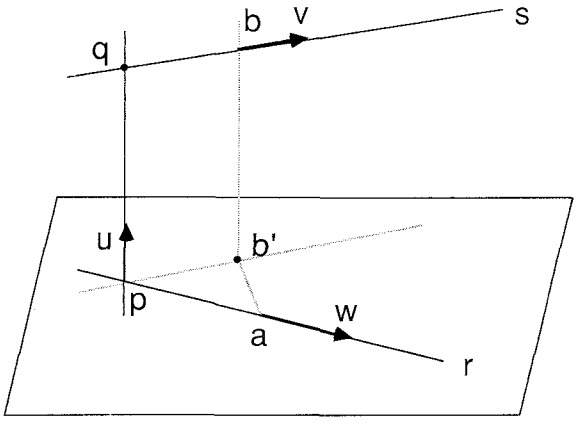
Si $\overrightarrow{pq} = ku$, $\mathrm{d}(p, q) = |k| = \left|\overrightarrow{pq} \cdot u\right|$. De fet, aquest producte escalar no depèn dels punts p i q sobre les rectes r i s; en efecte, de $\overrightarrow{ap} \cdot u = 0$ i $\overrightarrow{qb} \cdot u = 0$ resulta que

$$\overrightarrow{ab} \cdot u = (\overrightarrow{ap} + \overrightarrow{pq} + \overrightarrow{qb}) \cdot u = \overrightarrow{pq} \cdot u$$

i

$$\mathrm{d}(r, s) = \left|\overrightarrow{ab} \cdot u\right|,$$

on a i b són punts qualssevol de r i s respectivament, i u és un vector unitari perpendicular a les direccions de r i s.



Quan el sistema de referència $\{0;\ e_1, e_2, e_3\}$ és ortonormal, posant $u = \dfrac{w \wedge v}{\|w \wedge v\|}$, obtenim

$$\mathrm{d}(r, s) =$$

XIII.3 Isometries

Una aplicació $f : A_1 \longrightarrow A_2$ entre dos espais afins euclidians es diu una *isometria* si $\forall a, b \in A_1$

$$\mathrm{d}(f(a), f(b)) = \mathrm{d}(a, b).$$

Les isometries són sempre injectives, ja que si $f(a) = f(b)$,

$$0 = \mathrm{d}(f(a), f(b)) = \mathrm{d}(a, b)$$

i, per tant, $a = b$.

Les aplicacions que conserven l'estructura d'espai afí euclidià haurien d'ésser les afinitats tals que l'aplicació lineal associada conservés el producte escalar. La següent proposició ens assegura que aquestes apliacions són, precisament, les isometries.

Proposició 3.1 *Una aplicació $f : A_1 \longrightarrow A_2$ entre espais afins euclidans és una isometria si i només si és una afinitat i la seva aplicació lineal associada conserva el producte escalar.*

DEMOSTRACIÓ: Suposem que f és una isometria i definim una aplicació entre els espais vectorials associats de la següent manera: Fixat un punt $p \in A$,

$$\begin{array}{rccc} \phi: & E_1 & \longrightarrow & E_2 \\ & u = \overrightarrow{pa} & \longmapsto & \phi(u) = \overrightarrow{f(p)f(a)} \end{array}$$

(veure demostració de X.3.5). Aquesta aplicació conserva el producte escalar, ja que, donats $u = \overrightarrow{pa}$ i $v = \overrightarrow{pb}$,

$$\begin{aligned} \mathrm{d}(a,b)^2 = \|\overrightarrow{ab}\|^2 = \|\overrightarrow{ap} + \overrightarrow{pb}\|^2 = \|\overrightarrow{ap}\|^2 + \|\overrightarrow{pb}\|^2 + 2\overrightarrow{ap}\cdot\overrightarrow{pb} = \\ = \mathrm{d}(a,p)^2 + \mathrm{d}(p,b)^2 - 2u\cdot v. \end{aligned}$$

Anàlogament s'obté

$$\mathrm{d}(f(a),f(b))^2 = \mathrm{d}(f(a),f(p))^2 + \mathrm{d}(f(p),f(b))^2 - 2\,\overrightarrow{f(p)f(a)}\cdot\overrightarrow{f(p)f(b)};$$

d'on

$$u\cdot v = \overrightarrow{f(p)f(a)}\cdot\overrightarrow{f(p)f(b)} = \phi(u)\cdot\phi(v).$$

Per conservar el producte escalar ϕ és lineal (XII.1.1). Finalment,

$$\phi(\overrightarrow{ab}) = \phi(\overrightarrow{pb} - \overrightarrow{pa}) = \phi(\overrightarrow{pb}) - \phi(\overrightarrow{pa}) = \overrightarrow{f(p)f(b)} - \overrightarrow{f(p)f(a)} = \overrightarrow{f(a)f(b)}$$

i, per tant, f és una afinitat amb aplicació lineal associada ϕ.

Suposem ara que f és una afinitat i que $\tilde{f}$ conserva el producte escalar. Aleshores,

$$\mathrm{d}(f(a),f(b)) = \|\overrightarrow{f(a)f(b)}\| = \|\tilde{f}(\overrightarrow{ab})\| = \|\overrightarrow{ab}\| = \mathrm{d}(a,b)$$

i f és una isometria.□

Les isometries tenen, doncs, totes les propietats de les afinitats. En particular, conserven el paral·lelisme i la raó simple. A més a més, per (3.1) conserven també la perpendicularitat.

Dos espais afins euclidians A_1, A_2 es diuen *isomorfs* si existeix una isometria bijectiva $f : A_1 \longrightarrow A_2$. Escriurem llavors $A_1 \cong A_2$. Dos espais afins euclidians isomorfs són de la mateixa dimensió (X.1.6). Recíprocament, si $\dim A_1 = \dim A_2 = n < \infty$, $A_1 \cong A_2$. En efecte, siguin $\{e_1,\ldots,e_n\}$ i $\{u_1,\ldots,u_n\}$ bases ortonormals de E_1 i E_2 respectivament. Designem per $\phi : E_1 \longrightarrow E_2$ l'isomorfisme d'espais vectorials que aplica una base en l'altra: $\phi(e_i) = u_i$, $i = 1,\ldots,n$. ϕ conserva el producte escalar i qualsevol afinitat f amb aplicació lineal associada ϕ és un isomorfisme d'espais afins euclidians. Hem provat així el següent resultat:

Proposició 3.2 *Dos espais afins euclidians de dimensió finita són isomorfs si i només si tenen la mateixa dimensió.*□

Així, llevat d'isomorfismes, existeix un únic espai afí euclidià per a cada dimensió n. Un model particular n'és l'espai afí estàndard $\mathbf{R}^n$ amb el producte escalar que fa ortonormal la base $(1,0,\ldots,0),\ldots,(0,\ldots,0,1)$.

XIII.4 Classificació dels desplaçaments

Un *desplaçament* o *moviment* és una isometria d'un espai afí euclidià en ell mateix.

Dos desplaçaments $f, g : A \longrightarrow A$ són de la *mateixa classe* si existeix un isomorfisme d'espais afins euclidians $\phi : A \longrightarrow A$ que fa commutatiu el diagrama

$$\begin{array}{ccc} A & \xrightarrow{f} & A \\ \phi \downarrow & & \downarrow \phi \\ A & \xrightarrow{g} & A; \end{array}$$

és a dir, si $\phi \circ f = g \circ \phi$. Aquesta relació és clarament d'equivalència. Compareu-la amb la relació definida a (X.7) per a afinitats. La proposició següent correspon a la proposició (7.1) d'aquell capítol.

Proposició 4.1 *Dos desplaçaments $f, g : A \longrightarrow A$ són de la mateixa classe si i només si existeixen sistemes de referència ortonormals tals que les equacions de f en un coincideixen amb les equacions de g en l'altre.*

DEMOSTRACIÓ: Val la mateixa demostració de (X.7.1), observant simplement que $\tilde{\phi}$ conserva el producte escalar si i només si transforma bases ortonormals en bases ortonormals. □

Exemple:

Dues translacions, T_u i T_v, són de la mateixa classe si existeix un isomorfisme d'espais afins euclidians ϕ tal que $\forall a \quad \phi \circ T_u(a) = T_v \circ \phi(a)$. És a dir, $\phi(a) + \tilde{\phi}(u) = \phi(a) + v$. Per tant, $\tilde{\phi}(u) = v$. Això és possible si i només si u i v tenen la mateixa norma. Recordem, però, que com a afinitats i segons la classificació establerta a (X.7) T_u i T_v són de la mateixa classe sempre que $u \neq \vec{0} \neq v$.

Les equacions de T_u, $u \neq \vec{0}$, en una referència $\{p; \frac{u}{\|u\|}, e_2, \ldots, e_n\}$, són

$$\begin{cases} \bar{x}^1 = x^1 + \|u\| \\ \bar{x}^i = x^i \qquad i = 2, \ldots, n. \end{cases}$$

L'aplicació lineal associada a un desplaçament és una aplicació ortogonal; el seu determinant és, doncs, ± 1. Direm que un deplaçament $f : A \longrightarrow A$ és *propi* (o *directe*) si $\det \tilde{f} = +1$, i que és *impropi* (o *invers*)si $\det \tilde{f} = -1$.

XIII.5 Desplaçaments de la recta euclidiana

Sigui $f : A \longrightarrow A$ un desplaçament d'un espai afí euclidià A de dimensió 1. $\tilde{f}$ és ortogonal i, per tant, és la identitat I o $-I$. En el primer cas o $f = I$ ($\bar{x} = x$) o $f = T_u$ amb $u \neq \vec{0}$, i en un sistema de referencia com el de l'exemple de XIII.4

$$\bar{x} = x + \|u\|.$$

Si $\tilde{f} = -I$, f és una homotècia de raó -1 i, tal com vam veure a (X.2), té un únic punt fix. En qualsevol referència que tingui aquest punt com a origen, l'equació de f és

$$\bar{x} = -x.$$

f es diu llavors una *simetria central* de *centre* el punt fix.

XIII.6 Desplaçaments del pla euclidià

Sigui $f : A \longrightarrow A$ un desplaçament d'un espai afí euclidià A de dimensió 2. Si f és un desplaçament propi, $\tilde{f} \in SO(2)$ i en qualsevol base ortonormal $\{e_1, e_2\}$ la seva matriu és del tipus

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Si $\cos\theta \neq 1$, $\tilde{f}$ no té el valor propi 1 i, per tant, f té un únic punt fix q (X.6.1). Direm, aleshores, que f és un *gir* o *rotació d'angle* θ *i centre* q. Les seves equacions en la referència $\{q;\ e_1, e_2\}$ són

$$\begin{cases} \bar{x} = x\cos\theta - y\sin\theta \\ \bar{y} = x\sin\theta + y\cos\theta. \end{cases}$$

En el cas particular $\cos\theta = -1$, f és una *simetria central de centre* q.

Si $\cos\theta = 1$, $\tilde{f} = I$ i f és una translació de vector diferent de $\vec{0}$ o bé la identitat.

Quan f és un desplaçament impropi, $\tilde{f} \in O(2)$ amb $\det\tilde{f} = -1$. Existeix, aleshores, una base ortonormal $\{e_1, e_2\}$ en la qual la matriu de $\tilde{f}$ és

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Si f té un punt fix q, les seves equacions en la referència $\{q;\ e_1, e_2\}$ són

$$\begin{cases} \bar{x} = x \\ \bar{y} = -y \end{cases}$$

i es diu una *simetria axial*. La recta $q + \langle e_1 \rangle$ és de punts fixos i es diu l'*eix* de la simetria.

Si no existeix cap punt fix, les equacions de f en una referència $\{p;\ e_1, e_2\}$, amb $p \in A$ qualsevol, són

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = -y + d \end{cases}$$

amb $c \neq 0$. La recta $y = d/2$ és invariant. Si escollim l'origen sobre aquesta recta, $q = (x_0, d/2)$, obtenim $\overrightarrow{qf(q)} = (c, 0) = c\,e_1$. Les equacions de f en la referència $\{q;\ e_1, e_2\}$ són, per tant,

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = -y. \end{cases}$$

Aquest desplaçament és la *composició d'una simetria axial* d'eix $q + \langle e_1 \rangle$ *i una translació* paral·lela a l'eix, de vector $c\,e_1$. Aquesta composició és commutativa.

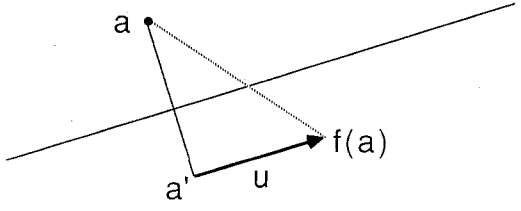
Observació:

Suposem donada una afinitat $f : A \longrightarrow A$, A espai afí euclidià de dimensió 2, en una referència no necessàriament ortonormal:

$$\bar{x} = Mx + b.$$

Llavors f és un desplaçament si i només si $\tilde{f}$ és ortogonal o equivalentment si $M^t G M = G$, on G és la matriu del producte escalar en aquesta referència. En cas d'ésser desplaçament, f és propi o impropi segons que $\det M$ sigui $+1$ o -1. En el primer cas, f és la identitat ($M = I$, $b = 0$), o una translació ($M = I$, $b \neq 0$), o un gir de centre l'únic punt fix i angle θ tal que $\text{traça}\, M = 2\cos\theta$.

Si f és un desplaçament impropi, es tracta d'una simetria axial si existeix una recta de punts fixos. Si no hi ha punts fixos, $f = T_u \circ s$, on s és una



simetria axial i T_u una translació paral·lela a l'eix de s. Aquest eix és l'única recta invariant per f i la seva direcció és el subespai E_1 de vectors propis de valor propi $+1$:

$$\text{Eix} = \{x \in A \mid \overrightarrow{xf(x)} \in E_1\}.$$

El vector u de la translació és precisament $u = \overrightarrow{xf(x)}$, per a qualsevol punt x de l'eix.

Exercici:

Si f és un desplaçament impropi, $\forall a \in A$ el punt mig del segment $\overline{af(a)}$ és de l'eix. Tenim així una altra manera de calcular l'eix.

Resum dels tipus de desplaçaments del pla euclidià

Sigui $\bar{x} = Mx + b$ un desplaçament del pla euclidià expressat en una referència qualsevol. Llavors:

$$\det M = 1 \quad (\cos\theta = \tfrac{1}{2}\,\text{traça}\,M) \begin{cases} \cos\theta = 1\ (M = I) & \begin{cases} b = 0 & \ldots. \text{ identitat} \\ b \neq 0 & \ldots. \text{ translació} \end{cases} \\ \cos\theta = -1\ (M = -I) & \ldots. \text{ simetria central} \\ |\cos\theta| \neq 1 & \ldots. \text{ rotació} \end{cases}$$

$$\det M = -1 \begin{cases} \nexists \text{ punts fixos} & \ldots. \text{ simetria axial composta amb una translació paral·lela a l'eix} \\ \exists \text{ punts fixos} & \ldots. \text{ simetria axial.} \end{cases}$$

XIII.7 Desplaçaments de l'espai euclidià tridimensional

Sigui $f : A \longrightarrow A$ un desplaçament d'un espai afí euclidià A de dimensió 3. Si f és propi, $\tilde{f} \in SO(3)$ i, en una base ortonormal convenient $\{e_1, e_2, e_3\}$, la seva matriu és de la forma

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si hi ha un punt fix q, hi ha tota una recta de punts fixos: $q + \langle e_3 \rangle$. El desplaçament es diu un *gir* o *rotació* d'angle θ i *eix* la recta de punts fixos. Les seves equacions en la referència $\{q;\ e_1, e_2, e_3\}$ són

$$\begin{cases} \bar{x} = x\cos\theta - y\sin\theta \\ \bar{y} = x\sin\theta + y\cos\theta \\ \bar{z} = z \end{cases}.$$

Si $\cos\theta = -1$ es diu que f és una *simetria axial.*

Si f no té cap punt fix, escollim de moment un origen p qualsevol. Siguin

$$\begin{cases} \bar{x} = x\cos\theta - y\sin\theta + c \\ \bar{y} = x\sin\theta + y\cos\theta + d \\ \bar{z} = z + n \end{cases}$$

les equacions de f en la referència $\{p;\ e_1, e_2, e_3\}$. Si $\cos\theta = 1$, f és una translació. Si $\cos\theta \neq 1$, en el sistema de punts fixos

$$\begin{cases} x(\cos\theta - 1) - y\sin\theta + c = 0 \\ x\sin\theta + y(\cos\theta - 1) + d = 0 \\ n = 0 \end{cases}$$

les dues primeres equacions tenen sempre una solució única (x_0, y_0). Com que estem suposant que no hi ha punts fixos, aquest sistema ha d'ésser incompatible, i per tant $n \neq 0$. La recta $x = x_0,\ y = y_0$ és invariant, i si escollim l'origen q sobre

ella, $q = (x_0, y_0, z)$, obtenim $\overrightarrow{qf(q)} = (0, 0, n)$. Les equacions de f en la referència $\{q;\ e_1, e_2, e_3\}$ són, per tant,

$$\begin{cases} \bar{x} = x\cos\theta - y\sin\theta \\ \bar{y} = x\sin\theta + y\cos\theta \\ \bar{z} = z + n \end{cases}$$

Això és una *rotació d'angle* θ i *eix* $q + \langle e_3 \rangle$ *composta amb una translació* de vector ne_3 paral·lel a l'eix. Aquesta composició és commutativa. Aquest tipus de desplaçament es diu *moviment helicoïdal.* Observem que la resta de desplaçaments propis són un cas particular de moviment helicoïdal per a $\cos\theta = 1$ i/o $n = 0$.

Si el desplaçament f és impropi, existeix una base ortonormal $\{e_1, e_2, e_3\}$ en la qual la matriu de $\vec{f}$ és

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Considerarem tres casos:

1. $\cos\theta = 1$ i existeix un punt fix q. Aleshores el pla $q + \langle e_1, e_2 \rangle$ és de punts fixos i f es diu una *simetria especular* respecte a aquest pla (*pla de simetria*). En la referència $\{q;\ e_1, e_2, e_3\}$ les equacions de f són

$$\begin{cases} \bar{x} = x, \\ \bar{y} = y \\ \bar{z} = -z. \end{cases}$$

2. $\cos\theta = 1$ i no hi ha cap punt fix. Les equacions de f en una referència $\{p;\ e_1, e_2, e_3\}$ són

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = y + d \\ \bar{z} = -z + n, \end{cases}$$

amb $c^2 + d^2 \neq 0$ perquè no hi hagi punts fixos. El pla $z = n/2$ és invariant i agafant l'origen sobre ell, $q = (x_0, y_0, n/2)$, tenim $\overrightarrow{qf(q)} = (c, d, 0)$. D'aquí resulta que les equacions de f en la referència $\{q;\ e_1, e_2, e_3\}$ són

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = y + d \\ \bar{z} = -z \end{cases}.$$

f és, per tant, la *composició d'una simetria especular* respecte a $q + \langle e_1, e_2 \rangle$ *i una translació de vector paral·lel al pla de simetria.* Aquesta composició és commutativa.

3. $\cos\theta \neq 1$. Llavors $\tilde{f}$ no té el valor propi 1 i, per tant, f té un únic punt fix q (X.6.1). En la referència $\{q;\ e_1, e_2, e_3\}$ les equacions de f són

$$\begin{cases} \bar{x} = x\cos\theta - y\sin\theta \\ \bar{y} = x\sin\theta + y\cos\theta \\ \bar{z} = -z \end{cases}$$

f és la *composició d'una simetria especular* respecte al pla $q + \langle e_1, e_2\rangle$ *i una rotació d'eix* $q + \langle e_3\rangle$ *perpendicular al pla de simetria.* Aquesta composició és commutativa. En el cas particular $\cos\theta = -1$, f es diu una *simetria central* de *centre* q.

Observació:

Suposem que A és un espai afí euclidià de dimensió 3 i $f : A \longrightarrow A$ una afinitat que en una referència no necessàriament ortonormal té les equacions

$$\bar{x} = Mx + b.$$

f és un desplaçament si i només si $M^tGM = G$, on G és la matriu del producte escalar en aquesta referència. Si f és desplaçament, és propi o impropi segons que $\det M$ sigui $+1$ o -1. En el primer cas f pot ésser una translació ($M = I$) o un gir, compost o no amb una translació. L'angle de gir θ compleix

$$2\cos\theta + 1 = \text{traça}\, M.$$

Si f és un gir i té punts fixos, l'eix és la recta de punts fixos. Si no hi ha punts fixos, l'eix és una recta invariant, de direcció el subespai E_1 de vectors propis de valor propi $+1$:

$$\text{Eix} = \{x \in A \mid \overrightarrow{xf(x)} \in E_1\}.$$

El vector de la translació és $u = \overrightarrow{xf(x)}$, on x és un punt qualsevol de l'eix.

En el segon cas f és una simetria especular, composta o no amb una translació de vector paral·lel al pla de simetria, o amb una rotació d'eix perpendicular a aquest pla.

El pla de simetria és un pla invariant. Per a tot $a \in A$, el punt mig del segment $\overline{af(a)}$ és d'aquest pla.

L'angle de la rotació ve donat per $2\cos\theta - 1 = \text{traça}\, M$.

L'eix és una recta invariant amb un punt fix, la direcció de la qual és al subespai E_{-1} de vectors propis de valor propi -1. Si aquest subespai no té dimensió 1, té dimensió 3 i f és una simetria central, de centre el punt fix.

Si $\cos\theta = 1$, es tracta d'una simetria especular seguida d'una translació i el vector de la translació és $u = \overrightarrow{xf(x)}$, on x és un punt qualsevol del pla de simetria.

Exercici:

Si f és una simetria axial, $\forall a \in A$ el punt mig del segment $\overline{af(a)}$ és de l'eix.

Resum dels tipus de desplaçaments de l'espai euclidià

Sigui $\bar{x} = Mx + b$ un desplaçament de l'espai euclidià tridimensional expressat en una referència qualsevol. Llavors:

$$\det M = 1 \quad (\cos\theta = \tfrac{1}{2}(\text{traça}\, M - 1)) \begin{cases} \exists \text{ punts fixos} \begin{cases} \cos\theta = 1\ (M = I) & \text{... identitat} \\ \cos\theta = -1 & \text{... simetria axial} \\ |\cos\theta| \neq 1 & \text{... rotació} \end{cases} \\ \nexists \text{ punts fixos} \begin{cases} \cos\theta = 1 & \text{... translació} \\ \cos\theta \neq 1 & \text{... moviment helicoïdal} \end{cases} \end{cases}$$

$$\det M = -1 \quad (\cos\theta = \tfrac{1}{2}(\text{traça}\, M + 1)) \begin{cases} \exists \text{ punts fixos} \begin{cases} \cos\theta = 1 & \text{... simetria especular} \\ \cos\theta = -1\ (M = -I) & \text{... simetria central} \\ |\cos\theta| \neq 1 & \text{... simetria especular composta amb una rotació perpendicular al pla de } \textit{simetria}. \end{cases} \\ \nexists \text{ punts fixos} & \text{... simetria especular composta amb una translació paral·lela al pla de simetria.} \end{cases}$$

XIII.8 Semblances

Una *semblança* és una aplicació $f : A \longrightarrow A$ d'un espai afí euclidià en ell mateix, que compleix

$$\mathrm{d}(f(a), f(b)) = r\, \mathrm{d}(a, b) \qquad \forall a, b \in A,$$

on r és un real fix ($r > 0$) que es diu la *raó* de la semblança. En particular, els desplaçaments són semblances de raó 1. Com en el cas dels desplaçaments, les semblances són sempre injectives. També és fàcil veure que la composició de dues semblances de raons r, r' és una semblança de raó rr'.

Proposició 8.1 *Una aplicació $f : A \longrightarrow A$, on A és un espai afí euclidià, és una semblança si i només si f és una afinitat i $\tilde{f}$ descompon en producte d'una aplicació ortogonal i una homotècia vectorial.*

DEMOSTRACIÓ: Suposem que f és una semblança de raó r i fixem un punt $p \in A$ qualsevol. Sigui h l'homotècia de centre p i raó r^{-1}.

$$\mathrm{d}(h(a),h(b)) = \|\overrightarrow{h(a)h(b)}\| = \|\tilde{h}(\overrightarrow{ab})\| = \|r^{-1}\overrightarrow{ab}\| = r^{-1}\mathrm{d}(a,b) \qquad \forall a,b$$

ens diu que h és una semblança de raó r^{-1} i, per tant, $g = h \circ f$ és una semblança de raó $r^{-1}r = 1$, és a dir, un desplaçament. Així doncs, $f = h^{-1} \circ g$ és una afinitat (producte de dues afinitats) i $\tilde{f} = \tilde{h}^{-1} \circ \tilde{g}$ amb $\tilde{g}$ ortogonal i $\tilde{h}^{-1} = rI$ una homotècia vectorial. Això demostra la primera part.

Observem que $g' = f \circ h$ també és un desplaçament i $f = g' \circ h^{-1}$. En general $g' \neq g$, però $\tilde{g}' = r^{-1}\tilde{f} = \tilde{g}$.

Suposem ara que f és una afinitat i $f = h \circ g$, amb h una homotècia de raó r i g un desplaçament. Per a tot parell a, b,

$$\mathrm{d}(f(a),f(b)) = \mathrm{d}(hg(a),hg(b)) = r\,\mathrm{d}(g(a),g(b)) = r\,\mathrm{d}(a,b).$$

Per tant, f és una semblança. Un raonament anàleg pot fer-se si $f = g \circ h$ amb g desplaçament i h homotècia. □

Proposició 8.2 *Tota semblança de raó $r \neq 1$ té un punt fix i només un.*

DEMOSTRACIÓ: Sigui $f = h \circ g$ una descomposició de la semblança com a producte d'una homotècia h de raó r i un desplaçament g. Per a tot vector u,

$$\tilde{f}(u) = r\tilde{g}(u),$$

d'on resulta que els únics valors propis possibles de $\tilde{f}$ són $\pm r \neq 1$. Llavors (X.6.1) assegura que existeix un únic punt fix. □

El punt fix q d'una semblança f de raó $r \neq 1$ es diu el *centre de la semblança.* Si en la primera part de la demostració de (8.1) agafem com a centre de l'homotècia h el punt q, els desplaçaments $g = h \circ f$ i $g' = f \circ h$ deixaran tots dos fix el punt q. Com que $\tilde{g} = \tilde{g}'$, resulta, doncs, que $g = g'$ i $f = h^{-1} \circ g = g \circ h^{-1}$.

Observació:

Considerem una afinitat $f : A \longrightarrow A$ d'equacions

$$\bar{x} = Mx + b$$

en una base no necessàriament ortonormal. f és una semblança de raó r si i només si $M = (rI)N$, N ortogonal, o, equivalentment, si $N^t G N = G$, $M = (rI)N$, on G és la matriu del producte escalar. És a dir, si i només si

$$M^t G M = r^2 G.$$

Aquesta igualtat ens permet comprovar si f és una semblança i calcular-ne la raó en cas afirmatiu. (Recordem que sempre $r > 0$).

Observem també que $\det M = r^n \det N = \pm r^n$, on n és la dimensió de l'espai A.

Una semblança f es diu *directa* si $\det \tilde{f} > 0$, i es diu *inversa* si $\det \tilde{f} < 0$.

Exercici:

Proveu que les úniques semblances de raó $\neq 1$ de la recta afí euclidiana són les homotècies.

XIII.9 Semblances de l'espai afí euclidià tridimensional

Sigui $f : A \longrightarrow A$ una semblança de raó $r \neq 1$, $\dim A = 3$. Denotem per q el centre de f. Llavors

$$f = h \circ g = g \circ h,$$

on g és un desplaçament que deixa fix q i h és l'homotècia de raó r i centre q.

Si f és una semblança directa, $\det \tilde{f} = r^3 \det \tilde{g} > 0$, d'on $\det \tilde{g} > 0$ i g és un desplaçament propi amb un punt fix q. g és, doncs, una rotació d'eix que passa per q.

Si f és una semblança inversa, $\det \tilde{f} = r^3 \det \tilde{g} < 0$, d'on $\det \tilde{g} < 0$ i g és un desplaçament impropi amb un punt fix. g és, doncs, una simetria especular seguida (o no) d'una rotació d'eix perpendicular al pla de simetria i intersecció el punt q. Podem, però, obtenir una interpretació geomètrica de f més clara de la forma següent. Considerem l'homotècia h' de centre g i raó $-r$. Tenim

$$h' = h \circ s,$$

on s és la simetria central de centre q. Per tant,

$$f = h \circ g = h \circ s \circ s \circ g = h' \circ g',$$

on $g' = s \circ g$ és un desplaçament propi amb un punt q fix. g' és, doncs, una rotació d'eix que passa per q i f s'obté component g' amb una homotècia de raó negativa $-r$. Hem demostrat així la següent proposició:

Proposició 9.1 *Les semblances de raó $r \neq 1$ de l'espai afí euclidià tridimensional són les rotacions seguides d'una homotècia de centre sobre l'eix de la rotació i raó $\pm r$.* □

XIII.10 Semblances del pla afí euclidià

Sigui $f : A \longrightarrow A$ una semblança de raó $r \neq 1$, $\dim A = 2$. Designem per q el centre de f. Aleshores,

$$f = h \circ g = g \circ h,$$

on g és un desplaçament que deixa fix q i h és l'homotècia de centre q i raó r.

Si f és directa, $\det \tilde{f} = r^2 \det \tilde{g} > 0$, d'on $\det \tilde{g} > 0$ i g és propi amb $g(q) = q$. g és, doncs, un gir de centre q.

Si f és inversa, $\det \tilde{f} = r^2 \det \tilde{g} < 0$, d'on $\det \tilde{g} < 0$ i g és impropi amb $g(q) = q$. g és una simetria axial respecte a un eix que passa per q.

Estudiarem ara les semblances en un model d'espai afí euclidià de dimensió 2 concret: els complexos. Aquest espai afí està format pel conjunt $A = \mathbf{C}$, l'espai vectorial de dimensió 2 sobre $\mathbf{R}$, $E = \mathbf{C}$, i l'aplicació

$$\begin{array}{rccc} \phi: & \mathbf{C}\times\mathbf{C} & \longrightarrow & \mathbf{C} \\ & (z_1, z_2) & \longmapsto & z_2 - z_1, \end{array}$$

on $z_j = a_j + ib_j \in \mathbf{C}$, $j = 1, 2$. Com a producte escalar a $\mathbf{C}$ considerarem el que fa que la base $\{1, i\}$ sigui ortonormal. És a dir,

$$\langle z_1, z_2 \rangle = a_1 a_2 + b_1 b_2.$$

En particular, $\|z_1\|^2 = {a_1}^2 + {b_1}^2$.

En el conjunt $\mathbf{C}$ tenim, a més a més d'aquestes estructures, un producte i una noció de conjugat d'un $z = a + ib \in \mathbf{C}$ que escriurem $\bar{z} = a - ib$. En particular, $\|z\|^2 = z \cdot \bar{z}$.

Una aplicació $f : \mathbf{C} \longrightarrow \mathbf{C}$ és una semblança de raó $r > 0$ si i només si és una afinitat i $\tilde{f}$ en la base $\{1, i\}$ té una matriu de la forma

$$\begin{pmatrix} ra & -rb \\ rb & ra \end{pmatrix}, \qquad a^2 + b^2 = 1, \qquad \text{si } f \text{ és directa;}$$

$$\begin{pmatrix} ra & rb \\ rb & -ra \end{pmatrix}, \qquad a^2 + b^2 = 1, \qquad \text{si } f \text{ és inversa.}$$

Donat $z = x + iy \in \mathbf{C}$, les seves coordenades en el sistema de referència ortonormal $\{0;\ 1, i\}$ són (x, y). Les equacions de f en el cas directe són, per tant, del tipus

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} ra & -rb \\ rb & ra \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} r(ax - by) + c \\ r(bx + ay) + d \end{pmatrix}.$$

Així doncs,

$$\begin{aligned} z^* = x^* + iy^* &= (r(ax - by) + c) + i(r(bx + ay) + d) = \\ &= r(a + ib)(x + iy) + c + id = \\ &= \alpha z + \beta, \end{aligned}$$

on $\alpha = r(a + ib)$, $\beta = (c + id)$. Observi's que $|\ \alpha\ | = r$.

Recíprocament, tota aplicació $f : \mathbf{C} \longrightarrow \mathbf{C}$ donada per $f(z) = z^* = \alpha z + \beta$, $\alpha, \beta \in \mathbf{C}$, és una semblança de raó el mòdul de α, $|\alpha|$. En efecte, si $\alpha = |\alpha|\,(a + ib)$, $\beta = c + id$, $z = x + iy$, $z^* = x^* + iy^*$, l'equació $z^* = \alpha z + \beta$ equival a

$$\begin{cases} x^* = |\alpha|\, ax - |\alpha|\, by + c \\ y^* = |\alpha|\, bx + |\alpha|\, ay + d \end{cases}$$

i això és, clarament, una semblança directa de raó $|\alpha|$.

Si f és inversa, les seves equacions en la referència $\{0;\ 1, i\}$ són del tipus

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} ra & rb \\ rb & -ra \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} r(ax+by)+c \\ r(bx-ay)+d \end{pmatrix}.$$

Així doncs,

$$\begin{aligned} z^* = x^* + iy &= (r(ax+by)+c) + i(r(bx-ay)+d) = \\ &= r(a+ib)(x-iy) + (c+id) = \\ &= \alpha\bar{z} + \beta, \end{aligned}$$

on $\alpha = r(a+ib)$, $\beta = c+id$, $|\alpha| = r$.

Recíprocament, donada $f : \mathbf{C} \longrightarrow \mathbf{C}$ per $f(z) = z^* = \alpha\bar{z} + \beta$, $\alpha, \beta \in \mathbf{C}$, si posem $\alpha = |\alpha|(a+ib)$, $\beta = c+id$, $z = x+iy$, $z^* = x^* + iy^*$, obtenim

$$\begin{cases} x^* = |\alpha|\, ax + |\alpha|\, by + c \\ y^* = |\alpha|\, bx - |\alpha|\, ay + d\,, \end{cases}$$

d'on resulta fàcilment que f és una semblança inversa de raó $|\alpha|$.

Observem que en aquestes expressions de les semblances del pla complex estan inclosos també els desplaçaments.

XIII.11 Alguns exemples i aplicacions

En aquest apartat anem a deduir uns quants resultats ben coneguts de la geometria elemental. El nostre objectiu és posar de manifest que la "geometria lineal" que hem estat estudiant és la "geometria ordinària" que ja coneixíem en part, encara que potser amb un altre llenguatge. Aquests exemples poden servir també de model per "traduir" altres resultats d'un llenguatge a l'altre.

Triangles

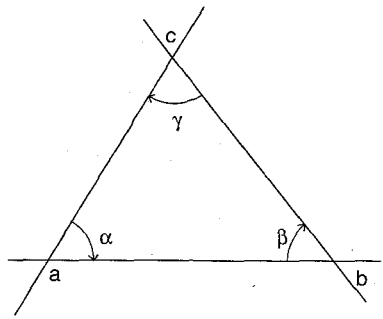
Situem-nos en el pla afí euclidià real.

Un *triangle* és un conjunt de tres punts linealment independents $\{a, b, c\}$ que anomenarem *vèrtexs*. Anomenarem *angles* d'un triangle $\{a, b, c\}$ els angles

$$\alpha = \widehat{\overrightarrow{ac}\,\overrightarrow{ab}}, \qquad \beta = \widehat{\overrightarrow{ba}\,\overrightarrow{bc}}, \qquad \gamma = \widehat{\overrightarrow{cb}\,\overrightarrow{ca}}.$$

La composició de les aplicacions ortogonals corresponents a aquests angles (XII.5) actua així:

$$\frac{\overrightarrow{ac}}{\|\overrightarrow{ac}\|} \overset{\alpha}{\longmapsto} \frac{\overrightarrow{ab}}{\|\overrightarrow{ab}\|} \overset{\beta}{\longmapsto} \frac{\overrightarrow{cb}}{\|\overrightarrow{cb}\|} \overset{\gamma}{\longmapsto} \frac{\overrightarrow{ca}}{\|\overrightarrow{ca}\|} = -\frac{\overrightarrow{ac}}{\|\overrightarrow{ac}\|}.$$



Aquesta composició és, per tant, $-I$ i

$$\alpha + \beta + \gamma = \pi.$$

Així doncs,

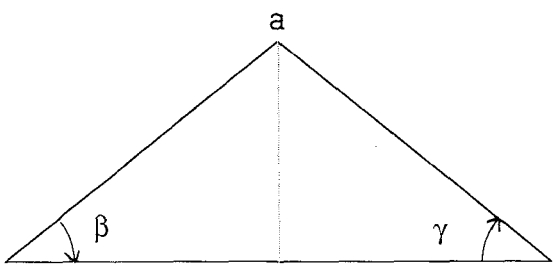
Proposició 11.1 *La suma dels angles d'un triangle és* π.□

Anomenarem *costats* d'un triangle a, b, c els segments determinats pels seus vèrtexs. Anomenarem *longitud* d'un costat la distància entre els vèrtexs corresponents.

Proposició 11.2 *Un triangle té dos angles iguals si i només si els dos costats oposats a aquestes angles tenen la mateixa longitud. Direm llavors que el triangle és* isòsceles.

DEMOSTRACIÓ: Suposem que $\{a, b, c\}$ és un triangle amb $\mathrm{d}(a, b) = \mathrm{d}(a, c)$. Designem per f la simetria axial d'eix $a + (\overrightarrow{ab} + \overrightarrow{ac})$. Llavors

$$\tilde{f}(\overrightarrow{ab} + \overrightarrow{ac}) = \overrightarrow{ab} + \overrightarrow{ac}.$$



A més a més,

$$(\overrightarrow{ab} - \overrightarrow{ac}) \cdot (\overrightarrow{ab} + \overrightarrow{ac}) = \|\overrightarrow{ab}\|^2 - \|\overrightarrow{ac}\|^2 = 0,$$

d'on $\tilde{f}(\overrightarrow{ab} - \overrightarrow{ac}) = -\overrightarrow{ab} + \overrightarrow{ac}$. Aquesta igualtat, juntament amb la de més amunt, dóna

$$\tilde{f}(\overrightarrow{ab}) = \overrightarrow{ac}\,, \quad \tilde{f}(\overrightarrow{ac}) = \overrightarrow{ab}\,, \quad \tilde{f}(\overrightarrow{cb}) = \overrightarrow{bc}.$$

Per tant, $\tilde{f}$ transforma l'angle $\beta = \widehat{\overrightarrow{ba}\,\overrightarrow{bc}}$ en l'angle $\widehat{\overrightarrow{ca}\,\overrightarrow{cb}} = -\gamma$. Llavors, per (XII.5.4), $\beta = \gamma$.

Suposem ara que $\beta = \gamma$. En una base ortonormal positiva (XII.4.2) tenim

$$\det\left(\frac{\overrightarrow{ba}}{\|\overrightarrow{ba}\|}\,,\ \frac{\overrightarrow{bc}}{\|\overrightarrow{bc}\|}\right) = \sin\beta = \sin\gamma = \det\left(\frac{\overrightarrow{cb}}{\|\overrightarrow{cb}\|}\,,\ \frac{\overrightarrow{ca}}{\|\overrightarrow{ca}\|}\right),$$

d'on

$$\det\left(\frac{\overrightarrow{ba}}{\|\overrightarrow{ba}\|} - \frac{\overrightarrow{ca}}{\|\overrightarrow{ca}\|}\,,\ \frac{\overrightarrow{bc}}{\|\overrightarrow{bc}\|}\right) = 0$$

i, per tant, existeix un k tal que

$$\frac{\overrightarrow{ba}}{\|\overrightarrow{ba}\|} - \frac{\overrightarrow{ca}}{\|\overrightarrow{ca}\|} = k\frac{\overrightarrow{bc}}{\|\overrightarrow{bc}\|} = \frac{k}{\|\overrightarrow{bc}\|}(\overrightarrow{ba} - \overrightarrow{ca}).$$

És a dir, $\|\overrightarrow{ba}\| = \|\overrightarrow{ca}\|$ i $\mathrm{d}(a,b) = \mathrm{d}(a,c)$. □

Punts notables del triangle

Es diuen *mitjanes* d'un triangle les rectes que passen per un vèrtex i el punt mig del costat oposat. Les tres mitjanes d'un triangle passen pel *baricentre* dels vèrtexs (Exercici IX.4).

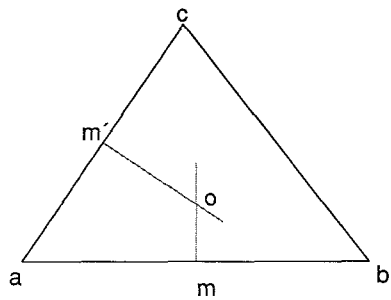
Es diuen *altures* d'un triangle les rectes que passen per un vèrtex i són perpendiculars a la recta determinada pels altres dos vèrtexs. En la situació de la proposició (11.2), la mitjana $a + \langle\overrightarrow{am}\rangle$, on $m = \frac{1}{2}b + \frac{1}{2}c$, és precisament l'altura per a, ja que

$$\overrightarrow{am}\cdot\overrightarrow{bc} = (\frac{1}{2}\overrightarrow{ab} + \frac{1}{2}\overrightarrow{ac})\cdot(\overrightarrow{ac} - \overrightarrow{ab}) = \frac{1}{2}(\|\overrightarrow{ac}\|^2 - \|\overrightarrow{ab}\|^2) = 0.$$

Les *mediatrius* d'un triangle $\{a,b,c\}$ són les rectes perpendiculars als costats que passen per llur punt mig. Siguin $m = \frac{1}{2}a + \frac{1}{2}b$, $m' = \frac{1}{2}a + \frac{1}{2}c$ i sigui O el punt d'intersecció de les mediatrius per m i m'. Observem que

$$\mathrm{d}(O,a)^2 = \|\overrightarrow{Oa}\|^2 = \|\overrightarrow{Om}\|^2 + \|\overrightarrow{ma}\|^2 = \|\overrightarrow{Om}\|^2 + \|\overrightarrow{mb}\|^2 = \|\overrightarrow{Ob}\|^2 = \mathrm{d}(O,b)^2.$$

Anàlogament, $\mathrm{d}(O,a) = \mathrm{d}(O,c)$. El triangle $\{O,b,c\}$ és, doncs, isòsceles i la seva altura talla $\overline{cb}$ en el punt mig. El punt O és, per tant, la intersecció de les tres mediatrius i es diu *circumcentre*. O és centre d'una circumferència que passa per a, b i c: la circumferència circumscrita al triangle.

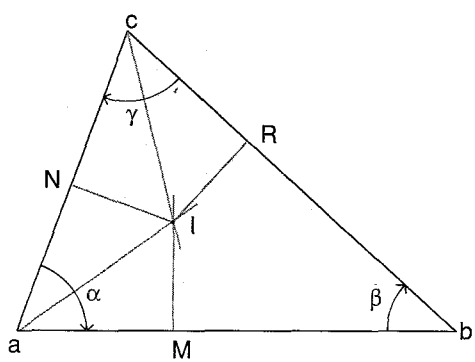


La *bisectriu* de l'angle α d'un triangle $\{a, b, c\}$ és la recta $a + \langle u \rangle$ tal que $\widehat{\overrightarrow{ac}u} = \widehat{u\overrightarrow{ab}}$.

Exercici:

Demostreu que en el triangle isòsceles de (11.2) la mitjana am és també la bisectriu de α.

Sigui I el punt d'intersecció de les bisectrius de α i de γ. Siguin M, N, R els peus de les perpendiculars per I a les rectes ab, ac, cb. I és de la bisectriu de β si



i només si $\mathrm{d}(I, R) = \mathrm{d}(I, M)$. Pel mateix motiu, com que I és de les bisectrius de α i γ, $\mathrm{d}(I, M) = \mathrm{d}(I, N) = \mathrm{d}(I, R)$. Per tant I és la intersecció de les bisectrius i es diu l'*incentre* del triangle. L'incentre és centre d'una circumferència tangent als costats del triangle a M, N i R: la circumferència inscrita.

Triangles semblants

Direm que dos triangles són *congruents* o *iguals* si existeix un desplaçament que transforma l'un en l'altre. Direm que són *semblants* si existeix una semblança que transforma l'un en l'altre.

Proposició 11.3 *Dos triangles són semblants si i només si es compleix una de les dues condicions equivalents:*

1. *Les longituds dels costats són proporcionals.*
2. *Els tres angles d'un triangle són iguals als angles de l'altre o són iguals a llurs oposats.*

DEMOSTRACIÓ: Les homotècies de raó positiva conserven els angles. Llavors (8.1) i (XII.5.4) ens diuen que les semblances directes conserven els angles i les inverses els inverteixen. Per tant, si dos triangles són semblants, es compleix 2. Per altra banda, en aplicar una semblança, les longituds dels costats queden multiplicades per la raó, i també es compleix 1.

Suposem ara que es compleix 1. Si $\{a,b,c\}$ i $\{a',b',c'\}$ són els triangles, posem

$$\begin{aligned} u &= \overrightarrow{ab} \quad & v &= \overrightarrow{ac} \quad & v-u &= \overrightarrow{bc} \\ u' &= \overrightarrow{a'b'} \quad & v' &= \overrightarrow{a'c'} \quad & v'-u' &= \overrightarrow{b'c'}. \end{aligned}$$

Per hipòtesi, $\|u\| = k\,\|u'\|$, $\|v\| = k\,\|v'\|$, $\|v-u\| = k\,\|v'-u'\|$. Designem per f l'afinitat que transforma $\{a,b,c\}$ en $\{a',b',c'\}$ i per h l'homotècia de raó k i centre un punt p qualsevol. Sigui $g = h \circ f$. Llavors

$$\tilde{g}(u) = ku', \quad \tilde{g}(v) = kv' \qquad \Rightarrow \qquad \|\tilde{g}(u)\| = \|u\|, \quad \|\tilde{g}(v)\| = \|v\|.$$

$$\begin{aligned} \tilde{g}(u)\cdot\tilde{g}(v) = k^2u'\cdot v' &= k^2\tfrac{1}{2}(\|u'\|^2 + \|v'\|^2 - \|u'-v'\|^2) = \\ &= \tfrac{1}{2}(\|u\|^2 + \|v\|^2 - \|u-v\|^2) = u\cdot v. \end{aligned}$$

$\tilde{g}$ és, doncs, ortogonal i g un desplaçament. Per tant, f és una semblança i $\{a,b,c\}$ i $\{a',b',c'\}$ són semblants.

Provem, per últim, que $2 \Rightarrow 1$. Amb les notacions d'abans,

$$\begin{aligned} \widehat{uv} = \pm\widehat{u'v'} \quad &\Rightarrow \quad \det\left(\frac{u}{\|u\|}, \frac{v}{\|v\|}\right) = \pm\det\left(\frac{u'}{\|u'\|}, \frac{v'}{\|v'\|}\right) \quad \Rightarrow \\ &\Rightarrow \quad \det(u,v) = \pm kk'\det(u',v'), \end{aligned}$$

on $k = \dfrac{\|u\|}{\|u'\|}$, $k' = \dfrac{\|v\|}{\|v'\|}$. Anàlogament,

$$\begin{aligned} \widehat{u(u-v)} = \pm\widehat{u'(u'-v')} \Rightarrow \quad & \det(u,u-v) = \pm k'k''\det(u',u'-v') \\ \Leftrightarrow \quad & \det(u,v) = \pm k'k''\det(u',v'), \end{aligned}$$

on $k'' = \dfrac{\|u-v\|}{\|u'-v'\|}$. Les dues igualtats obtingudes impliquen que $k' = k''$. Pel mateix motiu, resulta que $k = k'$ i, per tant,

$$\frac{\|u\|}{\|u'\|} = \frac{\|v\|}{\|v'\|} = \frac{\|u-v\|}{\|u'-v'\|}. \square$$

Proposició 11.4 *Si dos triangles $\{a,b,c\}$, $\{a',b',c'\}$ tenen un angle igual llevat del signe, $\alpha = \pm\alpha'$, i els dos costats que el formen proporcionals, $\dfrac{d(a,b)}{d(a',b')} = \dfrac{d(a,c)}{d(a',c')}$, els dos triangles són semblants.*

DEMOSTRACIÓ: Sigui f l'afinitat que aplica $\{a,b,c\}$ en $\{a',b',c'\}$. Amb la notació de (11.3) tenim $\|u\| = k\|u'\|$, $\|v\| = k\|v'\|$. L'aplicació lineal $\phi = k\tilde{f}$ és ortogonal, ja que

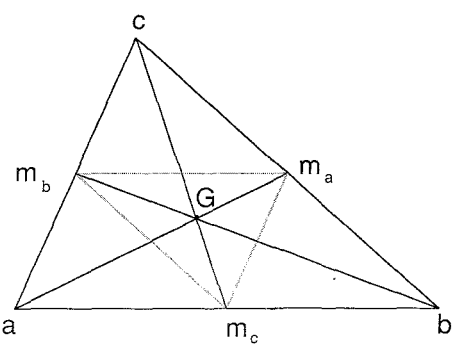
$$\|\phi(u)\| = \|ku'\| = \|u\| \qquad \|\phi(v)\| = \|kv'\| = \|v\|$$
$$\phi(u)\cdot\phi(v) = k^2u'\cdot v' = k^2\|u'\|\,\|v'\|\cos\widehat{u'v'} = \|u\|\,\|v\|\cos\widehat{uv} = u\cdot v.$$

Per tant, $\tilde{f} = k^{-1}\phi$ i f és una semblança.□

Considerem un triangle $\{a,b,c\}$ i els punts migs dels costats:

$$m_a = \frac{1}{2}b + \frac{1}{2}c \qquad m_b = \frac{1}{2}a + \frac{1}{2}c \qquad m_c = \frac{1}{2}a + \frac{1}{2}b.$$

L'homotècia de centre a i raó $\frac{1}{2}$ aplica $\{a,b,c\}$ en $\{a,m_c,m_b\}$. De $\overrightarrow{m_cm_b} = \overrightarrow{h(b)h(c)} = \tilde{h}(\overrightarrow{bc}) = \frac{1}{2}\overrightarrow{bc}$ es dedueix que les rectes m_cm_b i bc són paral·leles. Anàlogament, les rectes m_bm_a i m_am_c són paral·leles a ab i ca respectivament.



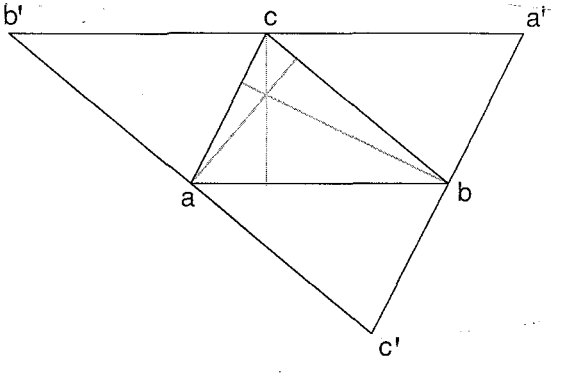
D'aquí resulta que els angles dels triangles $\{a,b,c\}$ i $\{m_a,m_b,m_c\}$ són iguals i, per (11.3), els triangles són semblants. Sigui f la semblança que passa de l'un a l'altre. Aquests dos triangles tenen les mateixes mitjanes i baricentre G, que serà un punt fix de f. A més a més, $\tilde{f}$ té vectors propis: $\overrightarrow{Ga}$, $\overrightarrow{Gb}$, $\overrightarrow{Gc}$; per tant, és una homotècia de raó $-\frac{1}{2}$, ja que $\|\overrightarrow{m_am_b}\| = \frac{1}{2}\|\overrightarrow{ab}\|$. En particular, obtenim

$$\mathrm{d}(G,m_a) = 2\mathrm{d}(G,a)$$

i igual per als altres vèrtexs.

Anem a fer servir la construcció anterior per demostrar que les altures d'un triangle es tallen en un punt: l'*ortocentre*.

Donat un triangle $\{a, b, c\}$, considerem les rectes que passen per un vèrtex i són paral·leles al costat oposat. Els punts d'intersecció d'aquestes rectes formen un triangle $\{a', b', c'\}$, els punts migs dels costats del qual són precisament a, b, c. (En efecte, siguin $m_{a'}, m_{b'}, m_{c'}$ els punts migs. Llavors les rectes ab i $m_{a'}m_{b'}$ són paral·leles perquè totes dues ho són a $\overline{a'b'}$. El mateix passa amb els altres dos parells. Llavors $m_{a'} \in \overline{ac'} \Leftrightarrow m_{b'} \in \overline{bc'} \Leftrightarrow m_{c'} \in \overline{cb'} \Leftrightarrow m_{a'} \in \overline{ab'}$. Per tant, $m_{a'} = a$, i anàlogament $m_{b'} = b$, $m_{c'} = c$).



Les altures de $\{a, b, c\}$ són, doncs, les mediatrius de $\{a', b', c'\}$ i, en particular, es tallen en un punt, com volíem provar.

Un triangle amb els tres costats iguals es diu *equilàter*. Per (11.2) els seus angles també són iguals: $\alpha = \beta = \gamma$. De $\alpha + \beta + \gamma = 3\alpha = \pi$ resulta fàcilment que $\cos\alpha = \frac{1}{2}$ i $\sin\alpha = \pm\frac{\sqrt{3}}{2}$ (el signe depèn de l'orientació del pla).

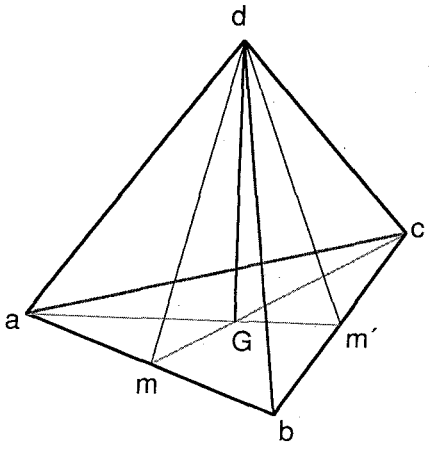
En un triangle equilàter les altures, mediatrius, mitjanes i bisectrius coincideixen i els seus punts d'intersecció també.

El tetràedre

Situem-nos ara en l'espai afí euclidià real tridimensional.

Un *tetràedre* és un conjunt de quatre punts linealment independents. Si la distància entre el seus vèrtexs és constant es diu que el tetràedre és *regular*. Es diuen *cares* d'un tetràedre els seus subconjunts de tres elements. Si un tetràedre és regular, les seves cares són triangles equilàters.

Sigui $\{a, b, c, d\}$ un tetràedre regular. Si m és el punt mig de $\overline{ab}$, la recta dm és l'altura de la cara $\{a, b, d\}$ i, per tant, perpendicular a $\overline{ab}$. Anàlogament, $\overline{cm}$ és perpendicular a $\overline{ab}$, i així tot el pla dmc és perpendicular a $\overline{ab}$. De la mateixa manera, si m' és el punt mig de $\overline{bc}$, el pla $dm'a$ és perpendicular a $\overline{cb}$. La intersecció d'aquests dos plans és la recta que passa per d i el baricentre G del triangle $\{a, b, c\}$.



Per estar en els dos plans, $\overline{dG}$ és perpendicular al pla abc. La recta dG passa, clarament, pel baricentre de $\{a, b, c, d\}$ i, per tant, aquest és el punt d'intersecció de les altures del tetràedre.

Observem per últim que el triangle $\{d, m, c\}$ és isòsceles. La seva altura és la perpendicular comú als costats $\overline{ab}$ i $\overline{dc}$ i els talla en els punts migs.

XIII.12 Nota històrica

Segons les idees del "Programa d'Erlangen", la geometria mètrica és l'estudi dels invariants pel grup de les isometries; és a dir, és una subgeometria de la geometria afí. Aquest punt de vista que Felix Klein (1849-1925) va introduir el 1872 ha perdurat i és el que hem adoptat.

En els capítols de Geometria del seu *Introductio*, Leonhard Euler (1707-1783) es va ocupar de buscar les corbes invariants per una isometria del pla, i els raonaments que desenvolupà porten a la conclusió que una tal isometria és una translació, o una rotació, o una translació seguida per una simetria axial d'eix la direcció de la translació. El 1776 demostrà que tota semblança directa del pla té un punt fix. L'any anterior, en connexió amb els seus treballs de Mecànica, havia intentat demostrar que un desplaçament propi de l'espai tenia una recta fixa, buscant una demostració que l'endomorfisme ortogonal associat tenia el valor propi 1, però no ho havia aconseguit. Lexell ho demostrà un any més tard, però no és fins els treballs de Michel Charles (1793-1880) el 1830 on es troba finalment una exposició completa i coherent de la classificació dels desplaçaments de l'espai i de llurs composicions.

XIII.13 Exercicis

1. Donats tres punts linealment independents de $\mathbf{R}^3$, demostreu que la in-

tersecció dels plans que passen per un punt i són perpendiculars a la recta determinada pels altres dos és una recta perpendicular al pla que conté els tres punts.

2. En un cub d'aresta 1 calculeu l'angle que formen les diagonals de dues cares contigües que concorren en un mateix vèrtex.

3. Donats tres nombres complexos de mòdul 1 que sumin 0, demostreu que formen un triangle equilàter.

4. Un nombre complex i les seves arrels quadrades són vèrtexs d'un triangle equilàter. Determineu-ne l'àrea.

5. Es defineix la *raó doble* de quatre nombres complexos a_1, a_2, a_3, a_4 com el quocient, si existeix,

$$(a_1\, a_2\, a_3\, a_4) = \frac{(a_1\, a_3\, a_4)}{(a_2\, a_3\, a_4)} \in \mathbf{C}.$$

Demostreu que $(a_1\, a_2\, a_3\, a_4) \in \mathbf{R}$ si i només si els quatre punts estan en una circumferència.

6. Siguin $(a_1, b_1), (a_2, b_2), (a_3, b_3)$ tres punts linealment independents del pla en una referència ortonormal. Demostreu que l'àrea del triangle que formen és

$$A = \frac{1}{2} \begin{vmatrix} a_1 & b_1 & 1 \\ a_2 & b_2 & 1 \\ a_3 & b_3 & 1 \end{vmatrix}.$$

Enuncieu i demostreu una fórmula anàloga per al volum d'un tetràedre.

7. Si g és el baricentre d'un triangle de vèrtexs a, b, c, demostreu que els triangles abg, bcg i cag tenen la mateixa àrea.

Enuncieu i demostreu un exercici anàleg per a un tetràedre.

8. Sigui $\{p; e_1, e_2, e_3\}$ una referència ortonormal de l'espai afí euclidià (A, E). Designem per s la simetria respecte a l'eix $p + \langle (a, b, c) \rangle$ i per $\tilde{s}$ el seu endomorfisme associat.

a) Demostreu que, per a tot $v \in E$, $\tilde{s}(v) + v$ és un vector propi de valor propi $+1$ (o bé és $\vec{0}$).

b) Deduïu de a) la matriu de $\tilde{s}$ en funció de a, b, c.

c) Trobeu les equacions de la simetria axial respecte a la recta

$$\begin{cases} 3x - 4y - 25 = 0 \\ z = 2 \end{cases}.$$

9. Sigui $\{p; e_1, e_2, e_3\}$ una referència ortonormal de l'espai afí euclidià (A, E), s la simetria especular respecte el pla $ax + by + cz + d = 0$ i $\tilde{s}$ l'endomorfisme associat a s.

a) Demostreu que, per a tot $v \in E$, $\tilde{s}(v) - v$ és ortogonal al pla de simetria.

b) Deduïu de a) la matriu de $\tilde{s}$ en funció de a, b, c.

c) Trobeu les equacions de la simetria especular respecte a $x+2y-3z+2 = 0$.

10. Lloc geomètric de les imatges del punt $(1, 1)$ per tots els girs de $\mathbf{R}^2$ d'angle $\frac{\pi}{2}$ i centre sobre la recta $x + y = 1$.

11. Donades dues rectes del pla no paral·leles, fem correspondre a cada punt M el punt mig M^* de les projeccions ortogonals de M sobre cada una de les rectes donades.

a) Estudieu la correspondència $M \longrightarrow M^*$.

b) Com han d'ésser les dues rectes donades perquè aquesta correspondència sigui una homotècia ?

12. Estudieu les semblances i els desplaçaments de la família d'afinitats

$$\begin{cases} \bar{x} &= (\frac{a}{3} + b)x + \frac{a}{3}y + \frac{a}{3}z + a \\ \bar{y} &= \frac{a}{3}x + (\frac{a}{3} + b)y + \frac{a}{3}z + a \\ \bar{z} &= \frac{a}{3}x + \frac{a}{3}y + (\frac{a}{3} + b)z + a. \end{cases}$$

13. Determineu el valor del paràmetre a perquè l'afinitat d'equacions

$$\begin{cases} \bar{x} &= ax - 20y - 15z + 46 \\ \bar{y} &= 20x + 9y - 12z + 16 \\ \bar{z} &= 15x - 12y + 16z - 60 \end{cases}$$

sigui una semblança.

Determineu-ne el centre, l'eix, l'angle i la raó.

14. Considerem la família d'afinitats de $\mathbf{R}^2$

$$\begin{cases} \bar{x} &= cy + a \\ \bar{y} &= x + b. \end{cases}$$

a) Determineu els valors dels paràmetres per als quals aquestes afinitats són desplaçaments i estudieu-los.

b) Per a quins valors dels paràmetres es compleix

$$\text{mín } \mathrm{d}(p, \bar{p}) = 4 \ ?$$

c) Determineu el lloc geomètric de les imatges d'un punt p donat per tots els desplaçaments de l'apartat b).

15. Pot ésser

$$\begin{pmatrix} 1 & -1 & -2 \\ 1 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

la matriu d'una semblança en alguna referència ? En cas afirmatiu, estudieu-la.

16. a) Demostreu que les simetries axials generen el grup dels desplaçaments del pla. En particular, tot gir g és producte de dues simetries axials, $g = s_1 \circ s_2$, on s_1 (o s_2) té per eix una recta prefixada que passa pel centre de g. Deduïu-ne un procediment geomètric per a trobar el centre del gir producte de dos donats.
b) Enuncieu i demostreu un exercici anàleg per a l'espai (substituint simetria axial per simetria especular, recta per pla i centre del gir per eix del gir).

17. Siguin $\{r_1, r_2\}, \{r_3, r_4\}, \{r_5, r_6\}$ els parells d'arestes oposades d'un tetràedre regular. Si s_i denota la simetria axial d'eix r_i, estudieu la composició $s_6 \circ s_5 \circ s_4 \circ s_3 \circ s_2 \circ s_1$.

18. Quines condicions ha de complir un quadrilàter del pla perquè existeixi una afinitat que el transformi en un quadrat ? Si n'existeix un en aquestes condicions, hi ha alguna homologia que el transformi en un quadrat ?

19. Estudieu el grup dels desplaçaments de $\mathbf{R}^2$ que deixen fix cada un dels següents conjunts:
i) un triangle equilàter;
ii) un quadrat;
iii) dues rectes que es tallen en un punt;
iv) una recta i un punt no contingut en ella.

20. Estudieu el grup dels desplaçaments de $\mathbf{R}^3$ que deixen fix cada un dels següents conjunts:
i) un tetràedre regular;
ii) dues rectes que es tallen en un punt;
iii) dues rectes que es creuen;
iv) un quadrat i un punt que no estan en un mateix pla;
v) un triangle equilàter i una recta que passa pel seu baricentre.

21. Sigui G el conjunt de totes les translacions i totes les homotècies.
a) Demostreu que G és un grup generat per les homotècies.
b) Demostreu que les translacions formen un subgrup normal de G.
c) Estudieu si el conjunt H de les homotècies de centre un punt p fixat és un subgrup normal de G.
d) Determineu el lloc geomètric dels centres de les homotècies dels conjunts

i) hHh^{-1}, on h és una homotècia donada;

ii) THT^{-1}, on T és una translació donada.

22. Demostreu que una afinitat és una semblança si i només si conserva l'ortogonalitat i si i només si conserva els angles.

XIII.14 Exercicis de programar

23. Feu un programa que, donades dues rectes de $\mathbf{R}^3$ que es creuin, trobi
a) les equacions de la perpendicular comuna (§1, exemple 2);
b) els peus d'aquesta perpendicular;
c) la distància entre les dues rectes (§2, II).

24. Prepareu un programa que, donat un hiperplà H de $\mathbf{R}^n$ i un punt p qualsevol, permeti calcular
a) la projecció ortogonal de p sobre H (§1, exemple 3);
b) la distància de p a H (§2, I).

25. Sigui $f : \mathbf{R}^3 \longrightarrow \mathbf{R}^3$ donat en la referència canònica per $f(x) = Mx + b$ amb $f \in O(3)$. Feu un programa que
a) calculi $\det M$ i $\operatorname{traça} M$;
b) decideixi si hi ha punts fixos (exercici X.22);
c) observant la taula de §7, classifiqui f;
d) calculi els elements de f que existeixin:

- pla de simetria;
- eix de rotació;
- vector de translació.

Nota: podeu utilitzar les matrius preparades a l'exercici XII.14 per provar el programa.

26. (Programa gràfic)

Feu un programa que permeti veure des de qualsevol angle una certa figura de l'espai. Podeu seguir el mètode següent:

a) Representeu la figura en una referència apropiada i guardeu en una matriu $A \in \mathbf{M}_{3\times n}(\mathbf{R})$ les coordenades dels seus punts.

b) Trieu la rotació que hi volgueu fer: l'angle φ i l'eix (X, Y o Z). Si R és la matriu d'aquesta rotació, calculeu RA.

c) Si hi voleu fer una translació de vector b, sumeu aquest vector a cada columna de la matriu. Podeu fer diverses rotacions i translacions successives repetint els mateixos passos.

d) Suposem que el pla de la pantalla és el YZ. Per a projectar-hi la matriu resultant cal fer el següent:

1) Eliminar la primera coordenada de tots els punts.
2) Per tal que el dibuix quedi centrat a la pantalla cal aplicar a cada punt (y_i, z_i) una afinitat

$$\begin{cases} \bar{y}_i & = & ry_i + c \\ \bar{z}_i & = & rz_i + d \end{cases}$$

on (c, d) és el centre de la pantalla segons les coordenades de l'ordinador i la raó r s'ha de triar de manera que la figura quedi dins dels límits d'aquestes coordenades.

e) Dibuixeu els punts $(\bar{y}_i, \bar{z}_i)$ obtinguts. Si a la figura inicial hi havia un segment que unia (x_i, y_i, z_i) amb (x_j, y_j, z_j), dibuixeu ara un segment que uneixi $(\bar{y}_i, \bar{z}_i)$ amb $(\bar{y}_j, \bar{z}_j)$.

Observació: aquest mateix programa serveix per a observar l'efecte d'una afinitat $f(x) = Mx + b$ sobre la figura. Només cal, en el segon pas, posar la matriu M en lloc de la R.